JC07 Rec'd PCT/PTO 04 JAN 2002

FORM PTO-1390 (Modified) (REV 11-2000) U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE **TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371**		ATTORNEY'S DOCKET NUMBER **MTR.0028US** U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 10/019965
INTERNATIONAL APPLICATION NO. **PCT/FR00/01908**	INTERNATIONAL FILING DATE **4 July 2000 (04.07.2000)**	PRIORITY DATE CLAIMED **5 July 1999 (05.07.1999)**

TITLE OF INVENTION

AUDIO ENCODING WITH HARMONIC COMPONENTS

APPLICANT(S) FOR DO/EO/US

FRANCOIS CAPMAN and CARLO MURGIA

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. ☒ This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. ☐ This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. ☒ This is an express request to begin national examination procedures (35 U.S.C. 371(f)). The submission must include itens (5), (6), (9) and (24) indicated below.
4. ☒ The US has been elected by the expiration of 19 months from the priority date (Article 31).
5. ☒ A copy of the International Application as filed (35 U.S.C. 371 (c) (2))
   - a. ☒ is attached hereto (required only if not communicated by the International Bureau).
   - b. ☐ has been communicated by the International Bureau.
   - c. ☐ is not required, as the application was filed in the United States Receiving Office (RO/US).
6. ☒ An English language translation of the International Application as filed (35 U.S.C. 371(c)(2)).
   - a. ☒ is attached hereto.
   - b. ☐ has been previously submitted under 35 U.S.C. 154(d)(4).
7. ☐ Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   - a. ☐ are attached hereto (required only if not communicated by the International Bureau).
   - b. ☐ have been communicated by the International Bureau.
   - c. ☐ have not been made; however, the time limit for making such amendments has NOT expired.
   - d. ☐ have not been made and will not be made.
8. ☒ An English language translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. ☒ An oath or declaration of the inventor(s) (35 U.S.C. 371 (c)(4)).
10. ☐ An English language translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).
11. ☒ A copy of the International Preliminary Examination Report (PCT/IPEA/409).
12. ☒ A copy of the International Search Report (PCT/ISA/210).

Items 13 to 20 below concern document(s) or information included:

13. ☐ An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
14. ☒ An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
15. ☒ A **FIRST** preliminary amendment.
16. ☐ A **SECOND** or **SUBSEQUENT** preliminary amendment.
17. ☐ A substitute specification.
18. ☐ A change of power of attorney and/or address letter.
19. ☐ A computer-readable form of the sequence listing in accordance with PCT Rule 13ter.2 and 35 U.S.C. 1.821 - 1.825.
20. ☐ A second copy of the published international application under 35 U.S.C. 154(d)(4).
21. ☐ A second copy of the English language translation of the international application under 35 U.S.C. 154(d)(4).
22. ☒ Certificate of Mailing by Express Mail
23. ☒ Other items or information:

    Thirteen (13) sheets of formal drawings.

PCTUS1/REV03

Express Mail No. EL857974142US

JC13 Rec'd PCT/PTO 04 JAN 2002

U.S. APPLICATION NO. (IF KNOWN, SEE 37 CFR 10/019965	INTERNATIONAL APPLICATION NO. PCT/FR00/01908	ATTORNEY'S DOCKET NUMBER MTR.0028US

24. The following fees are submitted:.

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

	CALCULATIONS	PTO USE ONLY
☐ Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$1040.00**		
☒ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$890.00**		
☐ International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO **$740.00**		
☐ International preliminary examination fee (37 CFR 1.482) paid to USPTO but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$710.00**		
☐ International preliminary examination fee (37 CFR 1.482) paid to USPTO and all claims satisfied provisions of PCT Article 33(1)-(4) **$100.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	**$890.00**	
Surcharge of **$130.00** for furnishing the oath or declaration later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (e)).	**$0.00**	

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE		
Total claims	42 - 20 =	22	x $18.00	**$396.00**	
Independent claims	2 - 3 =	0	x $84.00	**$0.00**	
Multiple Dependent Claims **(check if applicable).**			☐	**$0.00**	
TOTAL OF ABOVE CALCULATIONS =				**$1,286.00**	
☐ Applicant claims small entity status. See 37 CFR 1.27). The fees indicated above are reduced by 1/2.				**$0.00**	
SUBTOTAL =				**$1,286.00**	
Processing fee of **$130.00** for furnishing the English translation later than ☐ 20 ☐ 30 months from the earliest claimed priority date (37 CFR 1.492 (f)). +				**$0.00**	
TOTAL NATIONAL FEE =				**$1,286.00**	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31) **(check if applicable).** ☒				**$40.00**	
TOTAL FEES ENCLOSED =				**$1,326.00**	
				Amount to be: refunded	$
				charged	$

a. ☒ A check in the amount of **$1,326.00** to cover the above fees is enclosed.

b. ☐ Please charge my Deposit Account No. ________ in the amount of ________ to cover the above fees. A duplicate copy of this sheet is enclosed.

c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. **20-1504** A duplicate copy of this sheet is enclosed.

d. ☐ Fees are to be charged to a credit card. **WARNING:** Information on this form may become public. **Credit card information should not be included on this form.** Provide credit card information and authorization on PTO-2038.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Dan C. Hu
Trop, Pruner & Hu, P.C.
8554 Katy Freeway, Suite 100
Houston, Texas 77024

21906
PATENT TRADEMARK OFFICE

SIGNATURE

Dan C. Hu
NAME

40,025
REGISTRATION NUMBER

January 4, 2002
DATE

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants:	Francois Capman et al.	§	Group Art Unit:
Int'l Appl. No.:	PCT/FR00/01908	§	Examiner:
Int'l Filing Date:	4 July 2000	§	
For:	Audio Encoding with Harmonic Components	§	Atty. Dkt. No.: MTR.0028US

Box PCT
Commissioner for Patents
Washington DC 20231

PRELIMINARY AMENDMENT

Sir:

Prior to Examination, please amend the above-identified application as follows

In the Specification:

Page 1, at line 2, please insert the following paragraph:

--BACKGROUND OF THE INVENTION--

Page 2, at line 10, please insert the following paragraph:

--SUMMARY OF THE INVENTION--

Page 2, delete lines 29-32.

Page 2, at line 33, please insert the following paragraph:

--BRIEF DESCRIPTION OF THE DRAWINGS--

Page 4, at line 7, insert the following paragraph:

--DETAILED DESCRIPTION--

In the Abstract:

-- A fundamental frequency of the audio signal is estimated, and a spectrum of the audio signal is determined through a transform in the frequency domain of a frame of the audio signal. Data for coding a harmonic component of the audio signal, comprising data representative of spectral amplitudes associated with frequencies which are multiples of the fundamental frequency, are included in a digital output stream. The spectral amplitude

associated with one of the multiple frequencies is a local maximum of the modulus of the spectrum in the neighborhood of this multiple frequency. The data representative of spectral amplitudes associated with the multiple frequencies are obtained by means of cepstral coefficients calculated by transforming in the cepstral domain a compressed upper envelope of the spectrum of the audio signal.--

In the Claims:

Amend the following claims:

1. (Amended) A method of coding an audio signal, comprising the steps of:

estimating a fundamental frequency of the audio signal;

determining a spectrum of the audio signal through a transform into the frequency domain of a frame of the audio signal;

calculating cepstral coefficients by transforming in the cepstral domain a compressed upper envelope of the spectrum of the audio signal;

obtaining data representative of spectral amplitudes associated with frequencies multiple of the fundamental frequency by means of the calculated cepstral coefficients; and

including data for coding a harmonic component of the audio signal, comprising said data representative of the spectral amplitudes associated with frequencies multiple of the fundamental frequency, in a digital output stream,

wherein the spectral amplitude associated with one of said frequencies multiple of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency.

2. (Amended) The method as claimed in claim 1, further comprising the step of:

determining the compressed upper envelope by interpolation of said spectral amplitudes associated with the frequencies multiple of the fundamental frequency, with application of a spectral compression function.

3. (Amended) The method as claimed in claim 2, wherein the interpolation is performed between points each having a frequency multiple of the fundamental frequency as an abscissa and a spectral amplitude, compressed or uncompressed, associated with said multiple frequency as an ordinate.

4. (Amended) The method as claimed in claim 1, wherein the transformation in the cepstral domain of the compressed upper envelope is performed according to a nonlinear frequency scale.

5. (Amended) The method as claimed in claim 1, further comprising the steps of:

quantizing the cepstral coefficients to form said data representative of the spectral amplitudes associated with the frequencies multiple of the fundamental frequency.

6. (Amended) The method as claimed in claim 5, wherein the quantization of the cepstral coefficients is performed on a prediction residual for each of the cepstral coefficients.

7. (Amended) The method as claimed in claim 6, wherein the prediction residual for a cepstral coefficient is of the form $(cx[n,i] - \alpha(i)\ rcx\_q[n-1,i])/[2-\alpha(i)]$, where $cx[n,i]$ designates a current value of said cepstral coefficient, $rcx\_q[n-1,i]$ designates a previous value of the quantized prediction residual, and $\alpha(i)$ designates a prediction coefficient.

8. (Amended) The method as claimed in claim 6, further comprising the step of using different predictors to determine the prediction residuals for at least two of the cepstral coefficients.

9. (Amended) The method as claimed in claim 5, wherein the cepstral coefficients are distributed into several cepstral subvectors quantized separately by a vector quantization performed on a prediction residual of the cepstral coefficients.

10. (Amended) The method as claimed in claim 5, wherein the cepstral coefficients are normalized before quantization, by modifying the cepstral coefficient of order 0 so that the spectral amplitude associated with a frequency multiple of the fundamental frequency is represented exactly by the normalized cepstral coefficients.

11. (Amended) The method as claimed in claim 5, further comprising the step of:

transforming the cepstral coefficients by liftering in the cepstral domain prior to quantization.

12. (Amended) The method as claimed in claim 11, wherein the liftering is of the form $c_p(i) = [1+\gamma_2{}^i-\gamma_1{}^i].c(i)-(\mu^i/i)$, where $c_p(i)$ and $c(i)$ designate the cepstral coefficient of order $i>0$ respectively before and after liftering, γ_1 and γ_2 are coefficients lying between 0 and 1 and μ is a pre-emphasizing coefficient.

13. (Amended) The method as claimed in claim 12, wherein $\mu = (\gamma_2 - \gamma_1).c(1)$.

14. (Amended) The method as claimed in claim 11, further comprising the steps of :

recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients; and

adapting said liftering so as to minimize a discrepancy in modulus between the spectrum of the audio signal and at least one recalculated modulus value.

15. (Amended) The method as claimed in claim 11, further comprising the steps of:

recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients;

retransforming the cepstral coefficients by liftering and smoothing in the cepstral domain;

calculating minimum phases of the audio signal at frequencies multiple of the fundamental frequency on the basis of the retransformed cepstral coefficients; and

adapting the liftering performed prior to quantization so as to minimize a deviation between the spectrum of the audio signal and at least one complex value having a modulus value recalculated for a frequency multiple of the fundamental frequency and a phase value given by the minimum phase calculated for said multiple frequency.

16. (Amended) The method as claimed in claim 15, wherein the lifterings performed before and after quantization are adapted jointly so as to minimize said deviation,

and wherein parameters representative of the adapted liftering performed after quantization are included in the data for coding the harmonic component.

17. (Amended) The method as claimed in claim 14, wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in absolute value.

18. (Amended) The method as claimed in claim 14, further comprising the step of estimating a curve of spectral masking of the audio signal by means of a psycho-acoustic model, and wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve.

19. (Amended) The method as claimed in claim 1, wherein the spectrum of the audio signal and the cepstral coefficients resulting from the transformation of the compressed upper envelope are determined for successive mutually overlapping frames of N samples of the audio signal, and wherein said data representative of spectral amplitudes associated with the frequencies multiple of the estimated fundamental frequency, obtained by means of the cepstral coefficients calculated by transforming the compressed upper envelope, are included in the digital output stream for just one subset of the frames.

20. (Amended) The method as claimed in claim 19, wherein, for the frames which do not form part of said subset, data for quantizing an error of interpolation of the cepstral coefficients resulting from the transformation of the compressed upper envelope are included in the digital output stream.

21. (Amended) The method as claimed in claim 19, wherein, for the frames which do not form part of said subset, an optimal interpolator filter is determined for the cepstral coefficients resulting from the transformation of the compressed upper envelope and data representing said optimal interpolator filter are included in the digital output stream.

22. (Amended) An audio coder, comprising:

means for estimating a fundamental frequency of an audio signal;

means for determining a spectrum of the audio signal through a transform into the frequency domain of a frame of the audio signal;

means for calculating cepstral coefficients by transforming in the cepstral domain a compressed upper envelope of the spectrum of the audio signal;

means for obtaining data representative of spectral amplitudes associated with frequencies multiple of the fundamental frequency by means of the calculated cepstral coefficients; and

means for outputting a digital stream including data for coding a harmonic component of the audio signal,

wherein the data for coding a harmonic component of the audio signal include said data representative of spectral amplitudes associated with frequencies multiple of the fundamental frequency, and wherein the spectral amplitude associated with one of said frequencies multiple of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency.

23. (Amended) The audio coder as claimed in claim 22, further comprising:

means for determining the compressed upper envelope by interpolation of said spectral amplitudes associated with the frequencies multiple of the fundamental frequency, with application of a spectral compression function.

Add the following claims:

24. (New) The audio coder as claimed in claim 23, wherein the interpolation is performed between points each having a frequency multiple of the fundamental frequency as an abscissa and a spectral amplitude, compressed or uncompressed, associated with said multiple frequency as an ordinate.

25. (New) The audio coder as claimed in claim 22, wherein the transformation in the cepstral domain of the compressed upper envelope is performed according to a nonlinear frequency scale.

26. (New) The audio coder as claimed in claim 22, further comprising:

means for quantizing the cepstral coefficients to form said data representative of the spectral amplitudes associated with the frequencies multiple of the fundamental frequency.

27. (New) The audio coder as claimed in claim 26, wherein the quantization of the cepstral coefficients is performed on a prediction residual for each of the cepstral coefficients.

28. (New) The audio coder as claimed in claim 27, wherein the prediction residual for a cepstral coefficient is of the form (cx[n,i] - α(i) rcx_q[n-1,i])/[2-α(i)], where cx[n,i] designates a current value of said cepstral coefficient, rcx_q[n-1,i] designates a previous value of the quantized prediction residual, and α(i) designates a prediction coefficient.

29. (New) The audio coder as claimed in claim 27, further comprising a plurality of different predictors to determine the prediction residuals for at least two of the cepstral coefficients.

30. (New) The audio coder as claimed in claim 26, wherein the cepstral coefficients are distributed into several cepstral subvectors quantized separately by a vector quantization performed on a prediction residual of the cepstral coefficients.

31. (New) The audio coder as claimed in claim 26, wherein the cepstral coefficients are normalized before quantization, by modifying the cepstral coefficient of order 0 so that the spectral amplitude associated with a frequency multiple of the fundamental frequency is represented exactly by the normalized cepstral coefficients.

32. (New) The audio coder as claimed in claim 26, further comprising:

means for transforming the cepstral coefficients by liftering in the cepstral domain prior to quantization.

33. (New) The audio coder as claimed in claim 32, wherein the liftering is of the form $c_p(i) = [1+\gamma_2{}^i-\gamma_1{}^i].c(i)-(\mu^i/i)$, where $c_p(i)$ and c(i) designate the cepstral coefficient of order i>0 respectively before and after liftering, γ_1 and γ_2 are coefficients lying between 0 and 1 and μ is a pre-emphasizing coefficient.

34. (New) The audio coder as claimed in claim 33, wherein $\mu = (\gamma_2 - \gamma_1).c(1)$.

35. (New) The audio coder as claimed in claim 32, further comprising:

means for recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients; and

means for adapting said liftering so as to minimize a discrepancy in modulus between the spectrum of the audio signal and at least one recalculated modulus value.

36. (New) The audio coder as claimed in claim 32, further comprising:

means for recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients;

means for retransforming the cepstral coefficients by liftering and smoothing in the cepstral domain;

means for calculating minimum phases of the audio signal at frequencies multiple of the fundamental frequency on the basis of the retransformed cepstral coefficients; and

means for adapting the liftering performed prior to quantization so as to minimize a deviation between the spectrum of the audio signal and at least one complex value having a modulus value recalculated for a frequency multiple of the fundamental frequency and a phase value given by the minimum phase calculated for said multiple frequency.

37. (New) The audio coder as claimed in claim 35, wherein the lifterings performed before and after quantization are adapted jointly so as to minimize said deviation, and wherein parameters representative of the adapted liftering performed after quantization are included in the data for coding the harmonic component.

38. (New) The audio coder as claimed in claim 35, wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in absolute value.

39. (New) The audio coder as claimed in claim 35, further comprising means for estimating a curve of spectral masking of the audio signal by means of a psycho-acoustic model, and wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve.

40. (New) The audio coder as claimed in claim 22, wherein the spectrum of the audio signal and the cepstral coefficients resulting from the transformation of the compressed upper envelope are determined for successive mutually overlapping frames of N samples of the audio signal, and wherein said data representative of spectral amplitudes associated with the frequencies multiple of the estimated fundamental frequency, obtained by means of the cepstral coefficients calculated by transforming the compressed upper envelope, are included in the digital output stream for just one subset of the frames.

41. (New) The audio coder as claimed in claim 40, wherein, for the frames which do not form part of said subset, data for quantizing an error of interpolation of the cepstral coefficients resulting from the transformation of the compressed upper envelope are included in the digital output stream.

42. (New) The audio coder as claimed in claim 40, wherein, for the frames which do not form part of said subset, an optimal interpolator filter is determined for the cepstral coefficients resulting from the transformation of the compressed upper envelope and data representing said optimal interpolator filter are included in the digital output stream.

Remarks:

Allowance of all claims is respectfully requested. The Commissioner is authorized to charge any additional fees under 37 C.F.R. § 1.16 and § 1.17, or credit any overpayment to Deposit Account No. 20-1504 (MTR.0028US).

Date: 1/4/02

Respectfully submitted,

Dan C. Hu, Registration No. 40,025
TROP, PRUNER & HU, P.C.
8554 Katy Freeway, Suite 100
Houston, Texas 77024-1805
(713) 468-8880 [Phone]
(713) 468-8883 [Fax]

VERSIONS WITH MARKINGS TO SHOW CHANGES

IN THE CLAIMS:

New claims 24-42 have been added. Amendments of the claims are indicated below:

1. (Amended) A method of coding an audio signal, comprising the steps of:

[signal (x), in which] estimating a fundamental frequency [(F_0)] of the audio signal;

[signal is estimated,] determining a spectrum of the audio signal [is determined] through a transform into the frequency domain of a frame of the audio signal;

[signal, and] calculating cepstral coefficients by transforming in the cepstral domain a compressed upper envelope of the spectrum of the audio signal;

[data for coding a harmonic component of the audio signal, comprising] obtaining data representative of spectral amplitudes associated with frequencies [which are multiples] multiple of the fundamental frequency by means of the calculated cepstral coefficients; and

including data for coding a harmonic component of the audio signal, comprising said data representative of the spectral amplitudes associated with frequencies multiple of the fundamental frequency, [are included] in a digital output stream,

[stream (Φ), characterized in that] wherein the spectral amplitude associated with one of said frequencies [which are multiples] multiple of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency.

2. (Amended) The method as claimed in claim 1, further comprising the step of: [in which said data representative of spectral amplitudes associated with frequencies which are multiples of the fundamental frequency (F_0) are obtained by means of cepstral coefficients (cx_sup) calculated by transforming in the cepstral domain a compressed upper envelope (LX_sup) of the spectrum of the audio signal]

determining the compressed upper envelope by interpolation of said spectral amplitudes associated with the frequencies multiple of the fundamental frequency, with application of a spectral compression function.

3. (Amended) The method as claimed in claim 2, wherein the interpolation is performed between points each having a frequency multiple of the fundamental frequency as an abscissa and a spectral amplitude, compressed or uncompressed, associated with said multiple frequency as an ordinate [in which the compressed upper envelope (LX_sup) is

determined by interpolation of said spectral amplitudes associated with the frequencies which are multiples of the fundamental frequency (F_0), with application of a spectral compression function].

4. (Amended) The method as claimed in claim 1, wherein the transformation in the cepstral domain of the compressed upper envelope is performed according to a nonlinear frequency scale [3, in which the interpolation is performed between points whose abscissa is a frequency which is a multiple of the fundamental frequency (F_0) and whose ordinate is the spectral amplitude associated with said multiple frequency, compressed or uncompressed].

5. (Amended) The method as claimed in claim 1, further comprising the steps of: [any one of claims 2 to 4, in which the transformation in the cepstral domain of the compressed upper envelope (LX_sup) is performed according to a nonlinear frequency scale]

quantizing the cepstral coefficients to form said data representative of the spectral amplitudes associated with the frequencies multiple of the fundamental frequency.

6. (Amended) The method as claimed in claim 5, wherein the quantization of the cepstral coefficients is performed on a prediction residual for each of the cepstral coefficients [any one of claims 2 to 5, in which the cepstral coefficients (cx_sup) are quantized so as to form said data representative of the spectral amplitudes associated with the frequencies which are multiples of the fundamental frequency (F_0)].

7. (Amended) The method as claimed in claim 6, wherein the prediction residual for a cepstral coefficient is of the form (cx[n,i] - α(i) rcx_q[n-1,i])/[2-α(i)], where cx[n,i] designates a current value of said cepstral coefficient, rcx_q[n-1,i] designates a previous value of the quantized prediction residual, and α(i) designates a prediction coefficient [in which the quantization of the cepstral coefficients (cx_sup) pertains to a prediction residual for each of the cepstral coefficients].

8. (Amended) The method as claimed in claim 6, further comprising the step of using different predictors to determine the prediction residuals for at least two of the cepstral coefficients [7, in which the prediction residual for a cepstral coefficient is of the form (cx[n,i] - α(i) rcx_q[n-1,i])/[2-α(i)], where cx[n,i] designates a current value of said cepstral

coefficient, rcx_q[n-1,i] designates a previous value of the quantized prediction residual, and $\alpha(i)$ designates a prediction coefficient].

9. (Amended) The method as claimed in claim 5, wherein the cepstral coefficients are distributed into several cepstral subvectors quantized separately by a vector quantization performed on a prediction residual of the cepstral coefficients [7 or 8, in which different predictors are employed to determine the prediction residuals for at least two of the cepstral coefficients].

10. (Amended) The method as claimed in claim 5, wherein the cepstral coefficients are normalized before quantization, by modifying the cepstral coefficient of order 0 so that the spectral amplitude associated with a frequency multiple of the fundamental frequency is represented exactly by the normalized cepstral coefficients [any one of claims 6 to 9, in which the cepstral coefficients (cx_sup) are distributed into several cepstral subvectors quantized separately by a vector quantization pertaining to a prediction residual of the cepstral coefficients.].

11. (Amended) The method as claimed in claim 5, further comprising the step of:

transforming the cepstral coefficients by liftering in the cepstral domain prior to quantization [any one of claims 6 to 10, in which the cepstral coefficients (cx_sup) are normalized before quantization, by modifying the cepstral coefficient of order 0 in such a way that the spectral amplitude associated with a frequency which is a multiple of the fundamental frequency (F_0) is represented exactly by the normalized cepstral coefficients].

12. (Amended) The method as claimed in claim 11, wherein the liftering is of the form $c_p(i) = [1+\gamma_2{}^i-\gamma_1{}^i].c(i)-(\mu^i/i)$, where $c_p(i)$ and $c(i)$ designate the cepstral coefficient of order $i>0$ respectively before and after liftering, γ_1 and γ_2 are coefficients lying between 0 and 1 and μ is a pre-emphasizing coefficient [any one of claims 6 to 11, in which the cepstral coefficients (cx_sup) are transformed by liftering in the cepstral domain before being quantized].

13. (Amended) The method as claimed in claim 12, wherein $\mu = (\gamma_2 - \gamma_1).c(1)$. [in which the liftering is of the form $c_p(i) = [1 + \gamma_2^i - \gamma_1^i].c(i) - (\mu^i/_i)$, where $c_p(i)$ and $c(i)$ designate the cepstral coefficient of order i>0 respectively before and after liftering, γ_1 and γ_2 are coefficients lying between 0 and 1 and μ is a pre-emphasizing coefficient]

14. (Amended) The method as claimed in claim 11, further comprising the steps of: [13, in which $\mu = (\gamma_2 - \gamma_1).c(1)$]

recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients; and

adapting said liftering so as to minimize a discrepancy in modulus between the spectrum of the audio signal and at least one recalculated modulus value.

15. (Amended) The method as claimed in [any one of claims 12 to 14, in which a value of the modulus of the spectrum of the audio signal at at least one frequency which is a multiple of the fundamental frequency (F_0) is recalculated on the basis of the transformed and quantized cepstral coefficients (cx_sup_q), and said liftering is adapted in such a way as to minimize a discrepancy in modulus between the spectrum of the audio signal and at least one recalculated modulus value] claim 11, further comprising the steps of:

recalculating a value of the modulus of the spectrum of the audio signal at at least one frequency multiple of the fundamental frequency on the basis of the transformed and quantized cepstral coefficients;

retransforming the cepstral coefficients by liftering and smoothing in the cepstral domain;

calculating minimum phases of the audio signal at frequencies multiple of the fundamental frequency on the basis of the retransformed cepstral coefficients; and

adapting the liftering performed prior to quantization so as to minimize a deviation between the spectrum of the audio signal and at least one complex value having a modulus value recalculated for a frequency multiple of the fundamental frequency and a phase value given by the minimum phase calculated for said multiple frequency.

16. (Amended) The method as claimed in claim 15, wherein the lifterings performed before and after quantization are adapted jointly so as to minimize said deviation,

and wherein parameters representative of the adapted liftering performed after quantization are included in the data for coding the harmonic component [any one of claims 12 to 14, in which a value of the modulus of the spectrum of the audio signal at at least one frequency which is a multiple of the fundamental frequency (F_0) is recalculated on the basis of the transformed and quantized cepstral coefficients (cx_sup_q), the cepstral coefficients are retransformed by liftering and smoothing in the cepstral domain, minimum phases ($\varphi(k)$) of the audio signal at frequencies which are multiples of the fundamental frequency are calculated on the basis of the retransformed cepstral coefficients (cxl[n]), and the liftering performed before the quantization is adapted in such a way as to minimize a deviation between the spectrum of the audio signal and at least one complex value whose modulus has a value recalculated for a frequency which is a multiple of the fundamental frequency and whose phase is given by the minimum phase calculated for said multiple frequency].

17. (Amended) The method as claimed in claim 14, wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in absolute value [16, in which the lifterings performed before and after quantization are adapted jointly so as to minimize said discrepancy, and in which parameters (iLif) representative of the adapted liftering performed after quantization are included in the data for coding the harmonic component].

18. (Amended) The method as claimed in claim 14, further comprising the step of estimating a curve of spectral masking of the audio signal by means of a psycho-acoustic model, and wherein the minimized discrepancy for the adaptation of the liftering relates to at least one frequency multiple of the fundamental frequency, selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve [any one of claims 15 to 17, in which the minimized discrepancy for the adaptation of the liftering relates to at least one frequency which is a multiple of the fundamental frequency (F_0), selected on the basis of the magnitude of the modulus of the spectrum in absolute value].

19. (Amended) The method as claimed in claim 1, wherein the spectrum of the audio signal and the cepstral coefficients resulting from the transformation of the compressed upper envelope are determined for successive mutually overlapping frames of N samples of

the audio signal, and wherein said data representative of spectral amplitudes associated with the frequencies multiple of the estimated fundamental frequency, obtained by means of the cepstral coefficients calculated by transforming the compressed upper envelope, are included in the digital output stream for just one subset of the frames [any one of claims 15 to 17, in which a curve of spectral masking of the audio signal is estimated by means of a psycho-acoustic model, and the minimized discrepancy for the adaptation of the liftering relates to at least one frequency which is a multiple of the fundamental frequency (F_0), selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve].

20. (Amended) The method as claimed in claim 19, wherein, for the frames which do not form part of said subset, data for quantizing an error of interpolation of the cepstral coefficients resulting from the transformation of the compressed upper envelope are included in the digital output stream [2, in which the spectrum of the audio signal and the cepstral coefficients (cx_sup) resulting from the transformation of the compressed upper envelope are determined for successive frames of N samples of the audio signal which exhibit mutual overlaps, and in which said data representative of spectral amplitudes associated with the frequencies which are multiples of the estimated fundamental frequency (F_0), obtained by means of the cepstral coefficients calculated by transforming the compressed upper envelope, are included in the digital output stream (Φ) for just one subset of the frames].

21. (Amended) The method as claimed in claim 19, wherein, for the frames which do not form part of said subset, an optimal interpolator filter is determined for the cepstral coefficients resulting from the transformation of the compressed upper envelope and data representing said optimal interpolator filter are included in the digital output stream [20, in which, for the frames which do not form part of said subset, data (icx[n-1/2]) for quantizing an error (ecx[n-1/2]) of interpolation of the cepstral coefficients resulting from the transformation of the compressed upper envelope (LX_sup) are included in the digital output stream (Φ)].

22. (Amended) An audio coder, comprising:

means for estimating a fundamental frequency of an audio signal;

means for determining a spectrum of the audio signal through a transform into the frequency domain of a frame of the audio signal;

means for calculating cepstral coefficients by transforming in the cepstral domain a compressed upper envelope of the spectrum of the audio signal;

means for obtaining data representative of spectral amplitudes associated with frequencies multiple of the fundamental frequency by means of the calculated cepstral coefficients; and

means for outputting a digital stream including data for coding a harmonic component of the audio signal,

wherein the data for coding a harmonic component of the audio signal include said data representative of spectral amplitudes associated with frequencies multiple of the fundamental frequency, and wherein the spectral amplitude associated with one of said frequencies multiple of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency [The method as claimed in claim 20, in which, for the frames which do not form part of said subset, an optimal interpolator filter (128) is determined for the cepstral coefficients resulting from the transformation of the compressed upper envelope (LX_sup) and data (iP) representing said optimal interpolator filter are included in the digital output stream (Φ)].

23. (Amended) The audio coder as claimed in claim 22, further comprising: means for determining the compressed upper envelope by interpolation of said spectral amplitudes associated with the frequencies multiple of the fundamental frequency, with application of a spectral compression function [An audio coder, comprising means for executing a method according to any one of claims 1 to 22].

AUDIO CODING WITH HARMONIC COMPONENTS

The present invention relates to the field of the coding of audio signals. It applies in particular, but not exclusively, to the coding of speech, in narrowband or in broadband, in various coding bit rate ranges.

The design of an audio codec is aimed chiefly at providing a good compromise between the bit rate of the stream transmitted by the coder and the quality of the audio signal which the decoder is capable of reconstructing from this stream.

With this in mind, families of coders have in particular been developed which are based on analyzing the audio signal in the spectral domain: the coder estimates a fundamental frequency of the signal, representing its pitch, and the spectral analysis consists in determining parameters representing the harmonic structure of the signal at the frequencies which are integer multiples of this fundamental frequency. Modeling of the nonharmonic, or unvoiced, component may also be performed in the spectral domain. The parameters transmitted to the decoder typically represent the modulus of the spectrum of the voiced and unvoiced components. Added thereto is information representing either voiced/unvoiced decisions relating to various portions of the spectrum, or information regarding the probability of voicing of the signal, allowing the decoder to determine those portions of the spectrum in which it must use the voiced component or the unvoiced component.

These families of coders comprise the coders of the MBE type (standing for "Multi-Band Excitation"), or else the coders of the STC type (standing for "Sinusoidal Transform Coder"). By way of reference, mention may be made of US patents 4 856 068, 4 885 790, 4 937 873, 5 054 072, 5 081 681, 5 195 166, 5 216 747, 5 226 084,

5 226 108, 5 247 579, 5 473 727, 5 517 511, 5 630 011, 5 630 012, 5 649 050, 5 651 093, 5 664 051, 5 664 052, 5 684 926, 5 701 390, 5 715 365, 5 749 065, 5 752 222, 5 765 127, 5 774 837 and 5 890 108.

An aim of the present invention is to make it possible to improve the modeling of the modulus of the spectrum of the signal, in a coding scheme with analysis in the spectral domain.

The invention thus proposes a method of coding an audio signal, in which a fundamental frequency of the audio signal is estimated, a spectrum of the audio signal is determined through a transform in the frequency domain of a frame of the audio signal, and data for coding a harmonic component of the audio signal, comprising data representative of spectral amplitudes associated with frequencies which are multiples of the fundamental frequency, are included in a digital output stream. According to the invention, the spectral amplitude associated with one of said frequencies which are multiples of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency.

The invention also proposes an audio coder comprising means for implementing the above method.

Other features and advantages of the present invention will become apparent in the description below of non-limiting exemplary embodiments, with reference to the appended drawings, in which:

- figure 1 is a schematic diagram of an audio coder according to the invention;
- figures 2 and 3 are charts illustrating the formation of the audio signal frames in the coder of figure 1;
- figures 4 and 5 are graphs showing an exemplary

spectrum of the audio signal and illustrating the extraction of the upper and lower envelopes of this spectrum;

- figure 6 is a schematic diagram of an example of quantization means usable in the coder of figure 1;
- figure 7 is a schematic diagram of means usable to extract parameters relating to the phase of the nonharmonic component in a variant of the coder of figure 1;
- figure 8 is a schematic diagram of an audio decoder corresponding to the coder of figure 1;
- figure 9 is a flowchart of an exemplary procedure for smoothing spectral coefficients and for extracting minimum phases implemented in the decoder of figure 8;
- figure 10 is a schematic diagram of modules for analysis and for spectral mixing of harmonic and nonharmonic components of the audio signal;
- figures 11 to 13 are graphs showing examples of nonlinear functions usable in the analysis module of figure 10;
- figures 14 and 15 are charts illustrating a way of carrying out the temporal synthesis of the signal frames in the decoder of figure 8;
- figures 16 and 17 are graphs showing windowing functions usable in the synthesis of the frames according to figures 14 and 15;
- figures 18 and 19 are schematic diagrams of interpolation means usable in a variant embodiment of the coder and of the decoder;
- figure 20 is a schematic diagram of interpolation means usable in another variant embodiment of the coder;
- figures 21 and 22 are charts illustrating another way of carrying out the temporal synthesis of the signal frames in the decoder of figure 8, with the aid of an interpolation of parameters;
- figures 23 to 25 are schematic diagrams of variant

means of post-processing the cepstral coefficients representing the upper envelope of the spectrum of the signal in the coder of figure 1; and

- figure 26 is a partial schematic diagram of a decoder associated with a coder according to figure 25.

The coder and decoder described hereinbelow are digital circuits which can, as is customary in the field of audio signal processing, be embodied by programming a digital signal processor (DSP) or an application specific integrated circuit (ASIC).

The audio coder represented in figure 1 processes an audio input signal x which, in the nonlimiting example considered hereinbelow, is a speech signal. The signal x is available in digital form, for example at a sampling frequency F_e of 8 kHz. It is, for example, delivered by an analog/digital converter processing the amplified output signal from a microphone. The input signal x can also be formed from another version, analog or digital, coded or uncoded, of the speech signal.

The coder comprises a module 1 which forms successive frames of audio signal for the various processing operations performed, and an output multiplexer 6 which delivers an output stream Φ containing, for each frame, sets of quantization parameters from which a decoder will be capable of synthesizing a decoded version of the audio signal.

The structure of the frames is illustrated by figures 2 and 3. Each frame 2 is composed of a number N of consecutive samples of the audio signal x. The successive frames exhibit mutual time shifts corresponding to M samples, so that their overlap is L = N-M samples of the signal. In the example considered, where N = 256, M = 160 and L = 96, the

duration of the frames 2 is $N/F_e = 32$ ms, and a frame is formed every $M/F_e = 20$ ms.

In a conventional manner, the module 1 multiplies the samples of each frame 2 by a windowing function f_A, preferably chosen for its good spectral properties. The samples $x(i)$ of the frame being digitized from $i = 0$ to $i = N-1$, the analysis window $f_A(i)$ can thus be a Hamming window, expressed by:

$$f_A(i) = 0.54 + 0.46.\cos\left(2\pi \frac{i - (N-1)/2}{N}\right) \qquad (1)$$

or a Hanning window, expressed by:

$$f_A(i) = \frac{1}{2}\left(1 + \cos\left(2\pi \frac{i - (N-1)/2}{N}\right)\right) \qquad (2)$$

or else a Kaiser window, expressed by:

$$f_A(i) = \frac{I_0\left(\alpha\sqrt{1 - \left(\frac{i - (N-1)/2}{N}\right)^2}\right)}{I_0(\alpha)} \qquad (3)$$

where α is a coefficient equal, for example, to 6, and $I_0(.)$ designates the Bessel function of index 0.

The coder of figure 1 carries out an analysis of the audio signal in the spectral domain. It comprises a module 3 which calculates the fast Fourier transform (FFT) of each signal frame. The signal frame is shaped before being subjected to the FFT module 3: the module 1 appends N = 256 zero samples thereto so as to obtain the maximum resolution of the Fourier transform, and it moreover performs a circular permutation of the 2N = 512 samples so as to compensate for the phase effects resulting from the analysis window. This modification of the frame is illustrated by figure 3. The frame whose fast Fourier transform is calculated on 2N = 512 points commences with the last N/2 = 128 weighted samples of the frame, followed by the N = 256 zero samples, and terminates with the first N/2 = 128 weighted samples of the frame.

The FFT module 3 obtains the spectrum of the signal for each frame, whose modulus and phase are respectively denoted $|X|$ and φ_X, or $|X(i)|$ and $\varphi_X(i)$ for the frequency indices $i = 0$ to $i = 2N-1$ (by virtue of the symmetry of the Fourier transform and of the frames, we may confine ourselves to the values for $0 \leq i < N$).

A fundamental-frequency detector 4 estimates for each signal frame a value of the fundamental frequency F_0. The detector 4 can apply any known procedure for analyzing the speech signal of the frame to estimate the fundamental frequency F_0, for example a procedure based on the autocorrelation function or the AMDF function, possibly preceded by a module for whitening by linear prediction. The estimate can also be made in the spectral domain or in the cepstral domain. Another possibility is to evaluate the time intervals between the consecutive breaks in the speech signal which are attributable to closures of the talker's glottis occurring over the duration of the frame. Well-known procedures which can be used to detect such microbreaks are described in the following articles: M. Basseville et al., "Sequential detection of abrupt changes in spectral characteristics of digital signals" (IEEE Trans. on Information Theory, 1983, Vol. IT-29, No. 5, pages 708-723); R. Andre-Obrecht, "A new statistical approach for the automatic segmentation of continuous speech signals" (IEEE Trans. on Acous., Speech and Sig. Proc., Vol. 36, No. 1, January 1988); and C. MURGIA et al., "An algorithm for the estimation of glottal closure instants using the sequential detection of abrupt changes in speech signals" (Signal Processing VII, 1994, pages 1685-1688).

The estimated fundamental frequency F_0 forms the subject of a quantization, for example scalar, by a module 5, which provides the output multiplexer 6 with an index iF of quantization of the fundamental frequency for each frame of the signal.

The coder uses cepstral parametric modelings to represent an upper envelope and a lower envelope of the spectrum of the audio signal. The first step of the cepstral transformation consists in applying a spectral compression function to the modulus of the spectrum of the signal, which function may be a logarithmic or root function. The module 8 of the coder thus carries out, for each value $X(i)$ of the spectrum of the signal $(0 \leq i < N)$, the following transformation:

$$LX(i) = \mathrm{Log}(|X(i)|) \tag{4}$$

in the case of a logarithmic compression or

$$LX(i) = |X(i)|^{\gamma} \tag{5}$$

in the case of a root compression, γ being an exponent lying between 0 and 1.

The compressed spectrum LX of the audio signal is processed by a module 9 which extracts spectral amplitudes associated with the harmonics of the signal corresponding to the multiples of the estimated fundamental frequency F0. These amplitudes are then interpolated by a module 10 so as to obtain a compressed upper envelope denoted LX_sup.

It should be noted that the spectral compression could equivalently be performed after determining the amplitudes associated with the harmonics. It could also be performed after interpolation, and this would merely modify the form of the interpolation functions.

The module 9 for extracting the maxima takes account of any variation in the fundamental frequency over the analysis frame, errors which the detector 4 may make, as well as inaccuracies related to the discrete nature of the frequency sampling. To do this, the search for the amplitudes of the spectral peaks does not consist simply in taking the values $LX(i)$ corresponding to the indices i such that $i.F_e/2N$ is the frequency closest to a harmonic of frequency $k.F_0 (k \geq 1)$. The spectral

amplitude retained for a harmonic of order k is a local maximum of the modulus of the spectrum in the neighborhood of the frequency $k.F_0$ (this amplitude is obtained directly in compressed form when the spectral compression 8 is performed before the extraction of the maxima 9).

Figures 4 and 5 show an exemplary form of the compressed spectrum LX, where it may be seen that the maximum amplitudes of the harmonic peaks do not necessarily coincide with the amplitudes corresponding to the integer multiples of the estimated fundamental frequency F_0. Since the sides of the peaks are fairly steep, a small error in the positioning of the fundamental frequency F_0, amplified by the harmonic index k, may greatly distort the estimated upper envelope of the spectrum and cause poor modeling of the formant structure of the signal. For example, directly taking the spectral amplitude for the frequency $3.F_0$ in the case of figures 4 and 5 would produce a sizeable error in the extraction of the upper envelope in the neighborhood of the harmonic of order k = 3, although, in the example drawn, this relates to a zone of sizeable energy. By performing the interpolation on the basis of the actual maximum, this kind of error in estimating the upper envelope is avoided.

In the example represented in figure 4, the interpolation is performed between points whose abscissa is the frequency corresponding to the maximum of the amplitude of a spectral peak, and whose ordinate is this maximum, before or after compression.

The interpolation performed to calculate the upper envelope LX_sup is a simple linear interpolation. Of course, some other form of interpolation could be used (for example polynomial or spline).

In the preferred variant represented in figure 5, the

interpolation is performed between points whose abscissa is a frequency $k.F_0$ which is a multiple of the fundamental frequency (in fact the closest frequency in the discrete spectrum) and whose ordinate is the maximum amplitude, before or after compression, of the spectrum in the neighborhood of this multiple frequency.

By comparing figures 4 and 5, it may be seen that the mode of extraction according to figure 5, which repositions the peaks on the harmonic frequencies, leads to better accuracy with regard to the amplitude of the peaks which will be attributed by the decoder to the frequencies which are multiples of the fundamental frequency. A slight frequency displacement may occur in the position of these peaks, this not being very significant perceptually and anyway not being avoided either in the case of figure 4. In the case of figure 4, the anchoring points for the interpolation are one and the same as the vertices of the harmonic peaks. In the case of figure 5, these anchoring points must lie precisely at the frequencies which are multiples of the fundamental frequency, their amplitudes corresponding to those of the peaks.

The search interval for the amplitude maximum associated with a harmonic of rank k is centered on the index i of the frequency of the FFT closest to $k.F_0$, i.e. $i = \left\lfloor 2Nk\frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$, where $\lfloor a \rfloor$ designates the integer equal to or immediately less than the number a. The width of this search interval depends on the sampling frequency F_e, on the size 2N of the FFT and on the possible range of variation of the fundamental frequency. This width is typically of the order of some ten frequencies with the exemplary values considered earlier. It may be rendered adjustable as a function of the value F_0 of the fundamental frequency and of the

number k of the harmonic.

In order to improve the resolution in the low frequencies and hence to more faithfully represent the amplitudes of the harmonics in this zone, a nonlinear distortion of the frequency scale is carried out on the compressed upper envelope by a module 12 before the module 13 performs the inverse fast Fourier transform (IFFT) providing the cepstral coefficients cx_sup.

The nonlinear distortion allows more efficient minimization of the modeling error. It is, for example, performed on a frequency scale of Mel or Bark type. This distortion may possibly depend on the estimated fundamental frequency F_0. Figure 1 illustrates the case of the Mel scale. The relation between the frequencies F of the linear spectrum, expressed in hertz, and the frequencies F′ of the Mel scale is as follows:

$$F' = \frac{1000}{\log_{10}(2)} \times \log_{10}\left(1 + \frac{F}{1000}\right) \qquad (6)$$

In order to limit the transmission bit rate, a truncation of the cepstral coefficients cx_sup is performed. The IFFT module 13 need only calculate a cepstral vector of NCS cepstral coefficients of orders 0 to NCS-1. By way of example, NCS may be equal to 16.

Post-filtering in the cepstral domain, referred to as post-liftering, is applied by a module 15 to the compressed upper envelope LX_sup. This post-liftering corresponds to a manipulation of the cepstral coefficients cx_sup delivered by the IFFT module 13, which corresponds approximately to a post-filtering of the harmonic part of the signal by a transfer function having the conventional form:

$$H(z) = \left(1 - \mu z^{-1}\right)\frac{A(z/\gamma_1)}{A(z/\gamma_2)} \qquad (7)$$

where A(z) is the transfer function of a filter for linear prediction of the audio signal, γ_1 and γ_2 are

coefficients lying between 0 and 1, and μ is a pre-emphasizing coefficient, possibly zero. The relation between the post-liftered coefficient of order i, denoted $c_p(i)$, and the corresponding cepstral coefficient $c(i) = cx\_sup(i)$ delivered by the module 13 is then:

$$c_p(0) = c(0)$$

$$c_p(i) = \left(1 + \gamma_2^i - \gamma_1^i\right)c(i) - \frac{\mu^i}{i} \quad \text{for } i > 0 \tag{8}$$

The optional pre-emphasizing coefficient μ may be controlled by setting as constraint the preserving of the value of the cepstral coefficient cx_sup(1) relating to the slope. Specifically, the value of $c(1) = cx\_sup(1)$ of white noise filtered by the pre-emphasizing filter corresponds to the pre-emphasizing coefficient. The latter may thus be chosen as follows: $\mu = (\gamma_2 - \gamma_1).c(1)$.

After the post-lifter 15, a normalizing module 16 again modifies the cepstral coefficients by imposing the constraint of exact modeling of a point of the initial spectrum, which is preferably the point of greatest energy from among the spectral maxima extracted by the module 9. In practice, this normalization modifies only the value of the coefficient $c_p(0)$.

The normalizing module 16 operates as follows: it recalculates a value of the synthesized spectrum at the frequency of the maximum indicated by the module 9, by Fourier transform of the truncated and post-liftered cepstral coefficients, taking into account the nonlinear distortion of the frequency axis; it determines a normalizing gain g_N through the logarithmic difference between the value of the maximum as delivered by the module 9 and this value recalculated; and it adds the gain g_N to the post-liftered cepstral coefficient $c_p(0)$. This normalization may be viewed as being part of the post-liftering.

The post-liftered and normalized cepstral coefficients form the subject of a quantization by a module 18 which transmits corresponding quantization indices icxs to the output multiplexer 6 of the coder.

The module 18 can operate by vector quantization on the basis of cepstral vectors formed of post-liftered and normalized coefficients, here denoted cx[n] for the signal frame of rank n. By way of example, the cepstral vector cx[n] of NCS = 16 cepstral coefficients cx[n,0], cx[n,1], ..., cx[n,NCS-1] is distributed as four cepstral subvectors each containing four coefficients of consecutive orders. The cepstral vector cx[n] can be processed by the means represented in figure 6, forming part of the quantization module 18. These means implement, for each component cx[n,i], a predictor of the form:

$$cx_p[n,i] = (1-\alpha(i)).rcx[n,i] + \alpha(i).rcx[n-1,i] \quad (9)$$

where rcx[n] designates a residual prediction vector for the frame of rank n whose components are respectively denoted rcx[n,0], rcx[n,1], ..., rcx[n,NCS-1], and $\alpha(i)$ designates a prediction coefficient chosen so as to be representative of an assumed inter-frame correlation. After quantization of the residuals, this residual vector is defined by:

$$rcx[n,i] = \frac{cx[n,i] - \alpha(i).rcx\_q[n-1,i]}{2 - \alpha(i)} \quad (10)$$

where rcx_q[n-1] designates the quantized residual vector for the frame of rank n-1, whose components are respectively denoted rcx_q[n,0], rcx_q[n,1], ..., rcx_q[n,NCS-1].

The numerator of relation (10) is obtained by a subtractor 20, whose output vector components are divided by the quantities 2-α(i) at 21. For quantization purposes, the residual vector rcx[n] is subdivided into four subvectors, corresponding to the subdivision into four cepstral subvectors. On the basis

of a dictionary obtained by prior learning, the unit 22 undertakes the vector quantization of each subvector of the residual vector rcx[n]. This quantization can consist, for each subvector srcx[n], in selecting from the dictionary the quantized subvector srcx_q[n] which minimizes the quadratic error $\|srcx[n] - srcx\_q[n]\|^2$. The set icxs of quantization indices icx, corresponding to the addresses in the dictionary or dictionaries of the quantized residual subvectors srcx_q[n], is provided to the output multiplexer 6.

The unit 22 also delivers the values of the quantized residual subvectors, which form the vector rcx_q[n]. The latter is delayed by one frame at 23, and its components are multiplied by the coefficients $\alpha(i)$ at 24 so as to provide the vector to the negative input of the subtractor 20. The latter vector is, on the other hand, provided to an adder 25, the other input of which receives a vector formed by the components of the quantized residual rcx_q[n], respectively multiplied by the quantities $1-\alpha(i)$ at 26. The adder 25 thus delivers the quantized cepstral vector cx_q[n] which will be recovered by the decoder.

The prediction coefficient $\alpha(i)$ can be optimized separately for each of the cepstral coefficients. The quantization dictionaries may also be optimized separately for each four cepstral subvectors. Moreover, it is possible, in a manner known per se, to normalize the cepstral vectors before applying the prediction/quantization scheme, on the basis of the variance of the cepstra.

It should be noted that the above scheme for quantizing the cepstral coefficients cannot be applied other than in respect of certain of the frames only. For example, provision may be made for a second mode of quantization as well as a process for selecting that one of the two

modes which minimizes a least squares criterion with the cepstral coefficients to be quantized, and a bit indicating which of the two modes has been selected may be transmitted with the frame quantization indices.

The quantized cepstral coefficients cx_sup_q = cx_q[n] provided by the adder 25 are addressed to a module 28 which recalculates the spectral amplitudes associated with one or more of the harmonics of the fundamental frequency F_0 (figure 1). These spectral amplitudes are, for example, calculated in compressed form, by applying the Fourier transform to the quantized cepstral coefficients while taking account of the nonlinear distortion of the frequency scale used in the cepstral transformation. The amplitudes thus recalculated are provided to an adaptation module 29 which compares them with amplitudes of maxima determined by the extraction module 9.

The adaptation module 29 controls the post-lifter 15 in such a way as to minimize a discrepancy in modulus between the spectrum of the audio signal and the corresponding modulus values calculated at 28. This discrepancy in modulus can be expressed by a sum of absolute values of differences of amplitudes, compressed or otherwise, corresponding to one or more of the harmonic frequencies. This sum can be weighted as a function of the spectral amplitudes associated with these frequencies.

Optimally, the discrepancy in modulus taken into account in the adaptation of the post-liftering would take account of all the harmonics of the spectrum. However, in order to reduce the complexity of the optimization, the module 28 can resynthesize the spectral amplitudes for just one or more frequencies which are multiples of the fundamental frequency F_0 and which are selected on the basis of the magnitude of the modulus of the spectrum in absolute value. The

adaptation module 29 can, for example, consider the three most intense spectral peaks in the calculation of the discrepancy in modulus to be minimized.

In another embodiment, the adaptation module 29 estimates a curve of spectral masking of the audio signal by means of a psycho-acoustic model, and the frequencies taken into account in the calculation of the discrepancy in modulus to be minimized are selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve (it is, for example, possible to take the three frequencies for which the modulus of the spectrum most exceeds the masking curve). Various conventional methods can be used to calculate the masking curve from the audio signal. It is, for example, possible to use that developed by J.D. Johnston ("Transform Coding of Audio Signals Using Perceptual Noise Criteria", IEEE Journal on Selected Area in Communications, Vol. 6, No. 2, February 1988).

To carry out the adaptation of the post-liftering, the module 29 can use a filter identification model. A simpler method consists in predefining a collection of sets of post-liftering parameters, that is to say a collection of pairs γ_1, γ_2 in the case of post-liftering according to relations (8), in performing the operations incumbent on the modules 15, 16, 18 and 28 for each of these sets of parameters, and in retaining that of the sets of parameters which leads to the minimum discrepancy in modulus between the spectrum of the signal and the recalculated values. The quantization indices provided by the module 18 are then those which relate to the best set of parameters.

By a process similar to that for extracting the coefficients cx_sup representing the compressed upper envelope LX_sup of the spectrum of the signal, the coder determines the coefficients cx_inf representing a

compressed lower envelope LX_inf. A module 30 extracts from the compressed spectrum LX, spectral amplitudes associated with frequencies situated in zones of the spectrum which are intermediate with respect to the frequencies which are multiples of the estimated fundamental frequency F_0.

In the example illustrated by figures 4 and 5, each amplitude associated with a frequency situated in a zone intermediate between two successive harmonics $k.F_0$ and $(k+1).F_0$ corresponds simply to the modulus of the spectrum for the frequency $(k+1/2).F_0$ situated in the middle of the interval separating the two harmonics. In another embodiment, this amplitude could be an average of the modulus of the spectrum over a small span surrounding this frequency $(k+1/2).F_0$.

A module 31 carries out an interpolation, for example linear, of the spectral amplitudes associated with the frequencies situated in the intermediate zones so as to obtain the compressed lower envelope LX_inf.

The cepstral transformation applied to this compressed lower envelope LX_inf is performed according to a frequency scale resulting from a nonlinear distortion applied by a module 32. The IFFT module 33 calculates a cepstral vector of NCI cepstral coefficients cx_inf of orders 0 to NCI-1 representing the lower envelope. NCI is a number which may be substantially smaller than NCS, for example NCI = 4.

The nonlinear transformation of the frequency scale for the cepstral transformation of the lower envelope can be carried out to a scale which is finer at the high frequencies than at the low frequencies, thereby advantageously allowing good modeling of the unvoiced components of the signal at the high frequencies. However, to ensure homogeneity of representation between the upper envelope and the lower envelope, the

same scale will preferably be adopted in the module 32 as in the module 12 (Mel in the example considered).

The cepstral coefficients cx_inf representing the compressed lower envelope are quantized by a module 34, which may operate in the same manner as the module 18 for quantizing the cepstral coefficients representing the compressed upper envelope. In the case considered, where we restricted ourselves to NCI = 4 cepstral coefficients for the lower envelope, the vector thus formed is subjected to a prediction residual vector quantization performed by means identical to those represented in figure 6 but without subdivision into subvectors. The quantization index icx = icxi determined by the vector quantizer 22 for each frame relating to the coefficients cx_inf is provided to the output multiplexer 6 of the coder.

The coder represented in figure 1 does not comprise any particular device for coding the phases of the spectrum at the harmonics of the audio signal.

On the other hand, it comprises means 36-40 for coding time information related to the phase of the nonharmonic component represented by the lower envelope.

A spectral decompression module 36 and an IFFT module 37 form a temporal estimate of the frame of the non-harmonic component. The module 36 applies a decompression function which is the reciprocal of the compression function applied by the module 8 (that is to say an exponential or a $1/\gamma$ power function) to the compressed lower envelope LX_inf produced by the interpolation module 31. This provides the modulus of the estimated frame of the nonharmonic component, whose phase is taken equal to that φ_X of the spectrum of the signal X over the frame. The inverse Fourier transform performed by the module 37 provides the estimated frame

of the nonharmonic component.

The module 38 subdivides this estimated frame of the nonharmonic component into several time segments. The frame delivered by the module 37 being made up of 2N = 512 weighted samples, as illustrated by figure 3, the module 38 considers only the first N/2 = 128 samples and the last N/2 = 128 samples, and subdivides them, for example, into eight segments of 32 consecutive samples each representing 4 ms of signal.

For each segment, the module 38 calculates the energy equal to the sum of the squares of the samples, and forms a vector E1 formed of eight positive real components equal to the eight calculated energies. The largest of these eight energies, denoted EM, is also determined so as to be provided, with the vector E1, to a normalizing module 39. The latter divides each component of the vector E1 by EM, so that the normalized vector Emix is formed of eight components lying between 0 and 1. It is this normalized vector Emix, or weighting vector, which is subjected to the quantization by the module 40. The latter can carry out a vector quantization with a dictionary determined during prior learning. The quantization index iEm is provided by the module 40 to the output multiplexer 6 of the coder.

Figure 7 shows a variant embodiment of the means employed by the coder of figure 1 to determine the energy weighting vector Emix for the frame of the non-harmonic component. The spectral decompression and IFFT modules 36, 37 operate like those which bear the same references in figure 1. A selection module 42 is added so as to determine the value of the modulus of the spectrum subjected to the inverse Fourier transform 37. On the basis of the estimated fundamental frequency F_0, the module 42 identifies harmonic regions and non-harmonic regions of the spectrum of the audio signal.

For example, a frequency will be regarded as belonging to a harmonic region if it is located in a frequency interval centered on a harmonic $k.F_0$ and of width corresponding to a synthesized spectral line width, and to a nonharmonic region otherwise. In the nonharmonic regions, the complex signal subjected to the IFFT 37 is equal to the value of the spectrum, that is to say its modulus and its phase correspond to the values $|X|$ and φ_X provided by the FFT module 3. In the harmonic regions, this complex signal has the same phase φ_X as the spectrum and a modulus given by the lower envelope after spectral decompression 36. Proceeding thus according to figure 7 achieves more accurate modeling of the nonharmonic regions.

The decoder represented in figure 8 comprises an input demultiplexer 45 which extracts from the binary stream Φ, emanating from a coder according to figure 1, the quantization indices iF, icxs, icxi, iEm for the fundamental frequency F_0, the cepstral coefficients representing the compressed upper envelope, the coefficients representing the compressed lower envelope, and the weighting vector Emix, and distributes them respectively to modules 46, 47, 48 and 49. These modules 46-49 comprise quantization dictionaries similar to those of the modules 5, 18, 34 and 40 of figure 1, so as to restore the values of the quantized parameters. The modules 47 and 48 have dictionaries so as to form the quantized prediction residuals rcx_q[n], and they deduce therefrom the quantized cepstral vectors cx_q[n] with elements identical to the elements 23-26 of figure 6. These quantized cepstral vectors cx_q[n] provide the cepstral coefficients cx_sup_q and cx_inf_q processed by the decoder.

A module 51 calculates the fast Fourier transform of the cepstral coefficients cx_sup for each signal frame. The frequency scale of the compressed spectrum

resulting therefrom is modified nonlinearly by a module 52 applying the nonlinear transformation reciprocal to that of the module 12 of figure 1, and which provides the estimate LX_sup of the compressed upper envelope. A spectral decompression of LX_sup, carried out by a module 53, provides the upper envelope X_sup comprising the estimated values of the modulus of the spectrum at the frequencies which are multiples of the fundamental frequency F_0. The module 54 synthesizes the spectral estimate X_v of the harmonic component of the audio signal, through a sum of spectral lines centered on the frequencies which are multiples of the fundamental frequency F_0 and whose amplitudes (in modulus) are those given by the upper envelope X_sup.

Although the digital input stream Φ does not comprise any specific information regarding the phase of the spectrum of the signal at the harmonics of the fundamental frequency, the decoder of figure 8 is capable of extracting information regarding this phase from the cepstral coefficients cx_sup_q representing the compressed upper envelope. This phase information is used to assign a phase $\varphi(k)$ to each of the spectral lines determined by the module 54 in the estimate of the harmonic component of the signal.

As a first approximation, the speech signal may be regarded as being of minimum phase. Moreover, it is known that the minimum phase information may be deduced easily from cepstral modeling. This minimum phase information is therefore calculated for each harmonic frequency. The minimum phase assumption signifies that the energy of the synthesized signal is localized at the start of each period of the fundamental frequency F_0.

In order to be closer to a real speech signal, slight dispersion is introduced by means of a specific post-liftering of the cepstra during synthesis of the phase.

With this post-liftering, performed by the module 55 of figure 8, it is possible to emphasize the formant resonances of the envelope and hence to control the dispersion of the phases. This post-liftering is, for example, of the form (8).

To limit the phase breaks, it is preferable to smooth the post-liftered cepstral coefficients, this being performed by the module 56. The module 57 deduces from the post-liftered and smoothed cepstral coefficients the minimum phase assigned to each spectral line representing a harmonic peak of the spectrum.

The operations performed by the modules 56, 57 for smoothing and extracting the minimum phase are illustrated by the flowchart of figure 9. The module 56 examines the variations in the cepstral coefficients so as to apply lesser smoothing in the presence of abrupt variations than in the presence of slow variations. To do this, it performs the smoothing of the cepstral coefficients by means of a forget factor λ_c chosen as a function of a comparison between a threshold d_{th} and a distance d between two successive sets of post-liftered cepstral coefficients. The threshold d_{th} is itself adapted as a function of the variations of the cepstral coefficients.

The first step 60 consists in calculating the distance d between the two successive vectors relating to frames n-1 and n. These vectors, here denoted cxp[n-1] and cxp[n], correspond for each frame to the collection of NCS post-liftered cepstral coefficients representing the compressed upper envelope. The distance used may in particular be the Euclidean distance between the two vectors or else a quadratic distance.

Two smoothings are firstly performed, respectively by means of forget factors λ_{min} and λ_{max}, so as to determine a minimum distance d_{min} and a maximum distance d_{max}. The

threshold d_{th} is then determined in step 70 as being situated between the minimum and maximum distances d_{min}, d_{max}: $d_{th} = \beta.d_{max} + (1-\beta).d_{min}$, the coefficient β being, for example, equal to 0.5.

In the example represented, the forget factors λ_{min} and λ_{max} are themselves selected from among two distinct values, respectively λ_{min1}, λ_{min2} and λ_{max1}, λ_{max2} lying between 0 and 1, the indices λ_{min1}, λ_{max1} each being substantially nearer to 0 than the indices λ_{min2}, λ_{max2}. If $d > d_{min}$ (test 61), the forget factor λ_{min} is equal to λ_{min1} (step 62); otherwise, it is taken equal to λ_{min2} (step 63). In step 64, the minimum distance d_{min} is taken equal to $\lambda_{min}.d_{min} + (1-\lambda_{min}).d$. If $d > d_{max}$ (test 65), the forget factor λ_{max} is equal to λ_{max1} (step 66); otherwise, it is taken equal to λ_{max2} (step 67). In step 68, the minimum distance d_{max} is taken equal to $\lambda_{max}.d_{max} + (1-\lambda_{max}).d$.

If the distance d between the two consecutive cepstral vectors is greater than the threshold d_{th} (test 71), then a value λ_{c1} relatively close to 0 is adopted for the forget factor λ_c (step 72). In this case, the corresponding signal is regarded as being of nonstationary type, so that there is no need to keep a large memory of the earlier cepstral coefficients. If $d \leq d_{th}$, a value λ_{c2} which is not as close to 0 is adopted in step 73 for the forget factor λ_c, so as to further smooth the cepstral coefficients. The smoothing is performed in step 74, where the vector cxl[n] of smoothed coefficients for the current frame n is determined by:

$$cxl[n] = \lambda_c.cxl[n-1] + (1-\lambda_c).cxp[n] \qquad (11)$$

The module 57 then calculates the minimum phases $\varphi(k)$ associated with the harmonics $k.F_0$. In a known manner, the minimum phase for a harmonic of order k is given by:

$$\varphi(k) = -2.\sum_{m=1}^{NCS-1} cxl[n,m].\sin(2\pi mk F_0 / F_e) \qquad (12)$$

where cxl[n,m] designates the smoothed cepstral coefficient of order m for frame n.

In step 75, the harmonic index k is initialized to 1. To initialize the calculation of the minimum phase assigned to harmonic k, the phase $\varphi(k)$ and the cepstral index m are initialized to 0 and 1 respectively in step 76. In step 77, the module 57 adds the quantity $-2.cxl[n,m].\sin(2\pi mk.F_0/F_e)$ to the phase $\varphi(k)$. The cepstral index m is incremented in step 78 and compared with NCS in step 79. Steps 77 and 78 are repeated so long as m < NCS. When m = NCS, the calculation of the minimum phase is terminated for harmonic k, and the index k is incremented in step 80. The calculation of minimum phases 76-79 is rerun for the next harmonic so long as $k.F_0 < F_e/2$ (test 81).

In the exemplary embodiment according to figure 8, the module 54 takes account of a constant phase over the width of each spectral line, equal to the minimum phase $\varphi(k)$ provided for the corresponding harmonic k by the module 57.

The estimate X_v of the harmonic component is synthesized by summation of spectral lines positioned at the harmonic frequencies of the fundamental frequency F_0. During this synthesis, it is possible to position the spectral lines on the frequency axis with a higher resolution than the resolution of the Fourier transform. To do this, a reference spectral line is precalculated once and for all according to the higher resolution. This calculation can consist of a Fourier transform of the analysis window F_A with a transform size of 16 384 points, achieving a resolution of 0.5 Hz per point. The synthesis of each harmonic line is then performed by the module 54 by positioning on the

frequency axis the reference line with high resolution, and by undersampling this reference spectral line so as to reduce to the resolution of 16.625 Hz of the Fourier transform on 512 points. This enables the spectral line to be positioned accurately.

For the determination of the lower envelope, the FFT module 85 of the decoder of figure 8 receives the NCI quantized cepstral coefficients cx_inf_q of orders 0 to NCI - 1, and it advantageously supplements them with the NCS - NCI cepstral coefficients cx_sup_q of order NCI to NCS - 1 representing the upper envelope. Specifically, it may be estimated that, as a first approximation, the fast variations of the compressed lower envelope are well reproduced by those of the compressed upper envelope. In another embodiment, the FFT module 85 could consider only the NCI cepstral parameters cx_inf_q.

The module 86 converts the frequency scale in a manner reciprocal to the conversion carried out by the module 32 of the coder, so as to restore the estimate LX_inf of the compressed lower envelope, subjected to the spectral decompression module 87. At the output of the module 87, the decoder is furnished with a lower envelope X_inf comprising the values of the modulus of the spectrum in the valleys situated between the harmonic peaks.

This envelope X_inf will modulate the spectrum of a noise frame whose phase is processed as a function of the quantized weighting vector Emix extracted by the module 49. A generator 88 delivers a normalized noise frame whose 4-ms segments are weighted in a module 89 in accordance with the normalized components of the vector Emix provided by the module 49 for the current frame. This noise is white noise high-pass filtered so as to take account of the low level which in principle the unvoiced component has at the low frequencies. On

the basis of the energy-weighted noise, the module 90 forms frames of 2N = 512 samples by applying the analysis window f_A, the insertion of 256 zero samples and the circular permutation for phase compensation in accordance with what was explained with reference to figure 3. The Fourier transform of the resulting frame is calculated by the FFT module 91.

The spectral estimate X_{uv} of the nonharmonic component is determined by the spectral synthesis module 92 which performs a frequency-by-frequency weighting. This weighting consists in multiplying each complex spectral value provided by the FFT module 91 by the value of the lower envelope X_inf obtained for the same frequency by the spectral decompression module 87.

The spectral estimates X_v, X_{uv} of the harmonic (voiced in the case of a speech signal) and nonharmonic (or unvoiced) components are combined by a mixing module 95 controlled by a module 96 for analyzing the degree of harmonicity (or of voicing) of the signal.

The organization of these modules 95, 96 is illustrated by figure 10. The analysis module 96 comprises a unit 97 for estimating a frequency-dependent degree of voicing W from which are calculated four frequency-dependent gains, namely two gains g_v, g_{uv} controlling the relative magnitude of the harmonic and nonharmonic components in the synthesized signal, and two gains g_{v_φ}, g_{uv_φ} used to add noise to the phase of the harmonic component.

The degree of voicing W(i) is a continuously varying value lying between 0 and 1 determined for each frequency index i ($0 \le i < N$) as a function of the upper envelope X_sup(i) and of the lower envelope X_inf(i) which are obtained for this frequency i by the decompression modules 53, 87. The degree of voicing W(i) is estimated by the unit 97 for each frequency

index i corresponding to a harmonic of the fundamental frequency F_0, namely $i = \left\lfloor 2Nk\frac{F_0}{F_e} + \frac{1}{2} \right\rfloor$ for k = 1, 2, ..., by an increasing function of the ratio of the upper envelope X_sup to the lower envelope X_inf at this frequency, for example according to the formula:

$$W(i) = \min\left\{1, \frac{10.\log_{10}[X\_\sup(i)/X\_\inf(i)]}{Vth(F_0)}\right\} \qquad (13)$$

The threshold $Vth(F_0)$ corresponds to the average dynamic swing calculated over a purely voiced synthetic spectrum at the fundamental frequency. It is advantageously chosen to be dependent on the fundamental frequency F_0.

The degree of voicing W(i) for a frequency other than the harmonic frequencies is obtained simply as being equal to that estimated for the closest harmonic.

The gain $g_v(i)$, which depends on the frequency, is obtained by applying a nonlinear function to the degree of voicing W(i) (block 98). This nonlinear function has, for example, the form represented in figure 11:

$$g_v(i) = 0 \text{ if } 0 \le W(i) \le W1$$

$$g_v(i) = \frac{W(i) - W1}{W2 - W1} \text{ if } W1 < W(i) < W2 \qquad (14)$$

$$g_v(i) = 1 \text{ if } W2 \le W(i) \le 1$$

the thresholds W1, W2 being such that $0 < W1 < W2 < 1$. The gain g_{uv} can be calculated in a similar manner to the gain g_v (the sum of the two gains g_v, g_{uv} being constant, for example equal to 1), or deduced simply from the latter through the relation $g_{uv}(i) = 1 - g_v(i)$, as shown diagrammatically by the subtractor 99 in figure 10.

It is beneficial to be able to add noise to the phase of the harmonic component of the signal at a given frequency if the analysis of the degree of voicing shows that the signal is actually of nonharmonic type

at this frequency. To do this, the phase φ'_V of the mixed harmonic component is the result of a linear combination of the phases φ_v, φ_{uv} of the harmonic and nonharmonic components X_v, X_{uv} synthesized by the modules 54, 92.

The gains g_{v_φ}, g_{uv_φ} respectively applied to these phases are calculated from the degree of voicing W and also weighted as a function of the frequency index i, given that the adding of noise to the phase is actually useful only beyond a certain frequency.

A first gain g_{v1_φ} is calculated by applying a nonlinear function to the degree of voicing W(i), as shown diagrammatically by the block 100 in figure 10. This nonlinear function can have the form represented in figure 12:

$$g_{v1\_\varphi}(i) = G1 \quad \text{if } 0 \le W(i) \le W3$$

$$g_{v1\_\varphi}(i) = G1 + (1 - G1)\frac{W(i) - W3}{W4 - W3} \quad \text{if } W3 < W(i) < W4 \qquad (15)$$

$$g_{v1\_\varphi}(i) = 1 \quad \text{if } W4 \le W(i) \le 1$$

the thresholds W3 and W4 being such that $0 < W3 < W4 < 1$, and the minimum gain G1 lying between 0 and 1.

A multiplier 101 multiplies for each frequency of index i the gain g_{v1_φ} by another gain g_{v2_φ} dependent only on the frequency index i, so as to form the gain $g_{v\_\varphi}(i)$. The gain $g_{v2\_\varphi}(i)$ depends nonlinearly on the frequency index i, for example as indicated in figure 13:

$$g_{v2\_\varphi}(i) = 1 \quad \text{if } 0 \le i \le i1$$

$$g_{v2\_\varphi}(i) = 1 - (1 - G2)\frac{i - i1}{i2 - i1} \quad \text{if } i1 < i < i2 \qquad (16)$$

$$g_{v2\_\varphi}(i) = G2 \quad \text{if } i2 \le i \le 1$$

the indices i1 and i2 being such that $0 < i1 < i2 \le N$, and the minimum gain G2 lying between 0 and 1. The gain $g_{uv\_\varphi}(i)$ can be calculated simply as being equal to $1 - g_{v\_\varphi}(i) = 1 - g_{v1\_\varphi}(i).g_{v2\_\varphi}(i)$ (subtractor 102 of figure 10).

The complex spectrum Y of the synthesized signal is produced by the mixing module 95, which carries out the following mixing relation, for $0 \le i < N$:

$$Y(i) = g_v(i).|X_v(i)|.\exp[j\varphi'_V(i)]+g_{uv}(i).X_{uv}(i) \quad (17)$$

with $$\varphi'_V(i) = g_{v\_\varphi}(i).\varphi_v(i) + g_{uv\_\varphi}(i).\varphi_{uv}(i) \quad (18)$$

where $\varphi_v(i)$ designates the argument of the complex number $X_v(i)$ provided by the module 54 for the frequency of index i (block 104 of figure 10), and $\varphi_{uv}(i)$ desginates the argument of the complex number $X_{uv}(i)$ provided by the module 92 (block 105 of figure 10). This combination is carried out by the multipliers 106-110 and the adders 111-112 represented in figure 10.

The mixed spectrum Y(i) for $0 \le i < 2N$ (with $Y(2N-1-i) = Y(i)$) is then transformed into the time domain by the IFFT module 115 (figure 8). Only the first N/2 = 128 and the last N/2 = 128 samples of the frame of 2N = 512 samples produced by the module 115 are retained, and the circular permutation inverse to that illustrated by figure 3 is applied to obtain the synthesized frame of N = 256 samples weighted by the analysis window f_A.

The frames obtained successively in this manner are finally processed by the temporal synthesis module 116 which forms the decoded audio signal $\hat{x}$.

The temporal synthesis module 116 performs an overlap sum of frames modified with respect to those evaluated successively at the output of the module 115. The modification may be viewed in two steps illustrated by figures 14 and 15 respectively.

The first step (figure 14) consists in multiplying each frame 2′ delivered by the IFFT module 115 by a window $1/f_A$ inverse to the analysis window f_A employed by the module 1 of the coder. The samples of the frame 2″ resulting therefrom are therefore uniformly weighted.

The second step (figure 15) consists in multiplying the samples of this frame 2″ by a synthesis window f_S satisfying the following properties:

$$f_S(N-L+i) + f_S(i) = A \quad \text{for } 0 \le i < L \qquad (19)$$

$$f_S(i) = A \quad \text{for } L \le i < N-L \qquad (20)$$

where A designates an arbitrary positive constant, for example A = 1. The synthesis window $f_S(i)$ increases progressively from 0 to A for i going from 0 to L. It is, for example, a raised half-sinusoid:

$$f_S(i) = \frac{A}{2}.(1 - \cos[(i + 1/2)\pi / L]) \quad \text{for } 0 \le i < L \quad (21)$$

After having reweighted each frame 2″ by the synthesis window f_S, the module 116 positions the successive frames with their time shifts of M = 160 samples and their time overlaps of L = 96 samples, then it sums the frames thus positioned over time. Owing to the properties (19) and (20) of the synthesis window f_S, each sample of the decoded audio signal $\hat{x}$ thus obtained is assigned a uniform global weight, equal to A. This global weight originates from the contribution of a single frame if the sample has in this frame a rank i such that $L \le i < N - L$, and comprises the summed contributions of two successive frames if $0 \le i < L$ where $N - L \le i < N$.

It is thus possible to perform the temporal synthesis in a simple manner even if, as in the case considered, the overlap L between two successive frames is smaller than half the size N of these frames.

The two steps set forth above for modifying the signal frames may be merged into a single step. It is sufficient to precalculate a compound window $f_C(i) = f_S(i)/f_A(i)$ and simply to multiply the frames 2′ of N = 256 samples delivered by the module 115 by the compound window f_C before performing the overlap summation.

Figure 16 shows the shape of the compound window f_C in the case where the analysis window f_A is a Hamming window and the synthesis window f_S has the form given by relations (19) to (21).

Other forms of the synthesis window f_S satisfying relations (19) and (20) may be employed. In the variant of figure 17, it is a piecewise affine function defined by:

$$f_S(i) = A.i/L \text{ for } 0 \leq i < L \qquad (22)$$

In order to improve the quality of coding of the audio signal, the coder of figure 1 can increase the rate of formation and of analysis of the frames, so as to transmit more quantization parameters to the decoder. In the frame structure represented in figure 2, a frame of N = 256 samples (32 ms) is formed every 20 ms. These frames of 256 samples could be formed at a higher rate, for example 10 ms, two successive frames then having a shift of M/2 = 80 samples and an overlap of 176 samples.

Under these conditions, it is possible to transmit the complete sets of quantization parameters iF, icxs, icxi, iEm for just one subcollection of frames, and to transmit, for the other frames, parameters making it possible to perform a suitable interpolation at the level of the decoder. In the example envisaged hereinabove, the subcollection for which complete parameter sets are transmitted may consist of the frames of integer rank n, whose periodicity is M/F_e = 20 ms, and the frames for which an interpolation is performed may be those of half-integer rank n + 1/2 which are shifted by 10 ms with respect to the frames of the subcollection.

In the embodiment illustrated by figure 18, the notation cx_q[n-1] and cx_q[n] designates quantized

cepstral vectors determined, for two successive frames of integer rank, by the quantization module 18 and/or by the quantization module 34. These vectors comprise, for example, four consecutive cepstral coefficients each. They could also comprise more cepstral coefficients.

A module 120 performs an interpolation of these two cepstral vectors cx_q[n-1] and cx_q[n] so as to estimate an intermediate value cx_i[n-1/2]. The interpolation performed by the module 120 can be a simple arithmetic average of the vectors cx_q[n-1] and cx_q[n]. As a variant, the module 120 could apply a more sophisticated interpolation formula, for example polynomial, based also on the cepstral vectors obtained for frames earlier than frame n-1. Moreover, if more than one interpolated frame is interposed between two consecutive frames of integer rank, the interpolation takes account of the relative position of each interpolated frame.

With the aid of the means described above, the coder also calculates the cepstral coefficients cx[n-1/2] relating to the frame of half-integer rank. In the case of the upper envelope, these cepstral coefficients are those provided by the IFFT module 13 after post-liftering 15 (for example with the same post-liftering coefficients as for the previous frame n-1) and normalization 16. In the case of the lower envelope, the cepstral coefficients cx[n-1/2] are those delivered by the IFFT module 33.

A subtractor 121 forms the difference ecx[n-1/2] between the cepstral coefficients cx[n-1/2] calculated for the frame of half-integer rank and the coefficients cx_i[n-1/2] estimated by interpolation. This difference is provided to a quantization module 122 which addresses quantization indices icx[n-1/2] to the output multiplexer 6 of the coder. The module 122 operates,

for example, by vector quantization of the interpolation errors ecx[n-1/2] determined successively for the frames of half-integer rank.

This quantization of the interpolation error can be performed by the coder for each of the NCS + NCI cepstral coefficients used by the decoder, or for just some of them, typically those of smallest orders.

The corresponding means of the decoder are illustrated by figure 19. The decoder operates essentially like that described with reference to figure 8 to determine the signal frames of integer rank. An interpolation module 124 identical to the module 120 of the coder estimates the intermediate coefficients cx_i[n-1/2] from the quantized coefficients cx_q[n-1] and cx_q[n] provided by the module 47 and/or the module 48 from the indices icxs, icxi extracted from the stream Φ. A module for extracting parameters 125 receives the quantization index icx[n-1/2] from the input demultiplexer 45 of the decoder, and deduces therefrom the quantized interpolation error ecx_q[n-1/2] from the same quantization dictionary as that used by the module 122 of the coder. An adder 126 sums the cepstral vectors cx_i[n-1/2] and ecx_q[n-1/2] so as to provide the cepstral coefficients cx[n-1/2] which will be used by the decoder (modules 51-57, 95, 96, 115 and/or modules 85-87, 92, 95, 96, 115) so as to form the interpolated frame of rank n-1/2.

If just some of the cepstral coefficients have formed the subject of an interpolation error quantization, the others are determined by the decoder by a simple interpolation with no correction.

The decoder can also interpolate the other parameters F_0, Emix used to synthesize the signal frames. The fundamental frequency F_0 can be linearly interpolated, either in the time domain, or (preferably) directly in

the frequency domain. For the possible interpolation of the energy weighting vector Emix, it is appropriate to perform the interpolation after denormalization and while of course taking account of the time shifts between frames.

It should be noted that it is especially advantageous, in order to interpolate the representation of the spectral envelopes, to perform this interpolation in the cepstral domain. Unlike an interpolation performed on other parameters, such as the LSP coefficients (standing for "Line Spectrum Pairs"), the linear interpolation of the cepstral coefficients corresponds to the linear interpolation of the compressed spectral amplitudes.

In the variant represented in figure 20, the coder uses the cepstral vectors cx_q[n], cx_q[n-1], ..., cx_q[n-r] and cx_q[n-1/2] calculated for the last frames which have passed ($r \geq 1$) so as to identify an optimal interpolator filter which, when fed with the quantized cepstral vectors cx_q[n-r], ..., cx_q[n] relating to the frames of integer rank, delivers an interpolated cepstral vector cx_i[n-1/2] which exhibits a minimum distance with the vector cx[n-1/2] calculated for the last frame of half-integer rank.

In the example represented in figure 20, this interpolator filter 128 is present in the coder, and a subtractor 129 deducts its output cx_i[n-1/2] from the calculated cepstral vector cx[n-1/2]. A minimization module 130 determines the parameter set {P} of the interpolator filter 128, for which the interpolation error ecx[n-1/2] delivered by the subtractor 129 exhibits a minimum norm. This parameter set {P} is addressed to a quantization module 131 which provides a corresponding quantization index iP to the output multiplexer 6 of the coder.

As a function of the bit rate allocated in the stream Φ to the indices for quantizing the parameters {P} defining the optimal interpolator filter 128, it will be possible to adopt a finer or coarser quantization of these parameters, or a more or less elaborate form of the interpolator filter, or else to envisage several interpolator filters quantized differently for various vectors of cepstral coefficients.

In a simple embodiment, the interpolator filter 128 is linear, with r = 1:

$$cx\_i[n-1/2] = \rho.cx\_q[n-1] + (1-\rho).cx\_q[n] \quad (23)$$

and the parameter set {P} is limited to the coefficient ρ lying between 0 and 1.

From the indices iP for quantizing the parameters {P} obtained in the binary stream φ, the decoder reconstructs the interpolator filter 128 (to within quantization errors) and processes the spectral vectors cx_q[n-r], ..., cx_q[n] so as to estimate the cepstral coefficients cx[n-1/2] used to synthesize the frames of half-integer rank.

Generally, the decoder can use a simple interpolation method (without transmission of parameters by the coder for the frames of half-integer rank), and an interpolation method with incorporation of a quantized interpolation error (according to figures 17 and 18), or an interpolation method with an optimal interpolator filter (according to figure 19) to evaluate the frames of half-integer rank in addition to the frames of integer rank evaluated directly, as explained with reference to figures 8 to 13. The temporal synthesis module 116 can then combine the collection of these frames evaluated so as to form the synthesized signal $\hat{x}$ in the manner explained hereinbelow with reference to figures 14, 21 and 22.

As in the method of temporal synthesis described above, the module 116 performs an overlap sum of frames modified with respect to those evaluated successively at the output of the module 115, and this modification can be viewed in two steps of which the first is identical to that described above with reference to figure 14 (divide the samples of the frame 2′ by the analysis window f_A).

The second step (figure 21) consists in multiplying the samples of the renormalized frame 2″ by a synthesis window f_s' satisfying the following properties:

$$f_s'(i) = 0 \text{ for } 0 \leq i < N/2 - M/p \text{ and } N/2 + M/p \leq i < N \qquad (24)$$

$$f_s'(i) + f_s'(i + M/p) = A \text{ for } N/2 - M/p \leq i < N/2 \qquad (25)$$

where A designates an arbitrary positive constant, for example A = 1 and p is the integer such that the time shift between the successive frames (calculated directly and interpolated) is M/p samples, i.e. p = 2 in the example described. The synthesis window $f_s'(i)$ increases progressively for i going from N/2 - M/p to N/2. It is, for example, a raised sinusoid on the interval $N/2 - M/p \leq i < N/2 + M/p$. In particular, the synthesis window f_s' can, over this interval, be a Hamming window (as represented in figure 21) or a Hanning window.

Figure 21 shows the successive frames 2″ repositioned over time by the module 116. The hatching indicates the removed portions of the frames (synthesis window at 0). It may be seen that by performing the overlap sum of the samples of the successive frames, the property (25) ensures homogeneous weighting of the samples of the synthesized signal.

As in the method of synthesis illustrated by figures 14 and 15, the procedure for weighting the frames obtained by inverse Fourier transform of the spectra Y can be performed in a single step, with a compound window $f'_C(i) = f'_S(i)/f_A(i)$. Figure 22 shows the form of the compound window f'_C in the case where the windows f_A and f'_S are of Hamming type.

Like the method of temporal synthesis illustrated by figures 14 to 17, that illustrated by figures 14, 21 and 22 makes it possible to take into account an overlap L between two analysis frames (for which the analysis is performed completely) which is smaller than half the size N of these frames. In general, this latter method is applicable when the successive analysis frames exhibit mutual time shifts M of more than N/2 samples (possibly even of more than N samples if a very low bit rate is required), the interpolation leading to a collection of frames whose mutual time shifts are less than N/2 samples.

The interpolated frames can form the subject of a reduced transmission of coding parameters, as is described above, but this is not compulsory. This embodiment makes it possible to retain a relatively large interval M between two analysis frames, and hence to limit the transmission bit rate required, whilst limiting the discontinuities which are liable to appear by virtue of the size of this interval with respect to the typical timescales for the variations in the parameters of the audio signal, in particular the cepstral coefficients and the fundamental frequency.

Figures 23 to 25 show other embodiments of the means employed to process the cepstral coefficients cx_sup delivered by the IFFT module 13 of figure 1, representing the upper envelope.

In the three cases, the post-liftering module 15, normalizing module 16, quantization module 18 and module for calculating the spectral amplitudes 28 are essentially identical to those described previously with reference to figure 1. Furthermore, modules for post-liftering 140, for smoothing 141 and for extracting the minimum phase 142 are provided so as to process the post-liftered and quantized cepstral coefficients cx_sup_q delivered by the quantization module 18. These modules 140-142 operate essentially like the corresponding modules 55-57 of the decoder of figure 8.

In the embodiment shown in figure 23, the adaptation module 144 accomplishes a function similar to that of the module 29 of figure 1. However, the adaptation is not carried out solely on the basis of the modulus of the spectrum. The module 144 determines the best set of coefficients for the post-lifter 15 by minimizing the discrepancy between the spectrum of the audio signal, in terms of modulus $|X|$ and phase φ_X, and of the recalculated complex values for one or more of the harmonics of the fundamental frequency. The moduli of these latter complex values are given by the calculation module 28, and their phases correspond to the minimum phases $\varphi(k)$ provided by the extraction module 142. To carry out the adaptation, the module 144 can take into account any appropriate distance in the complex plane, for example the Euclidean distance.

Thus, the adaptation of the post-lifter 15 by the module 144 takes account in a combined manner of frequency aspects of the signal, which are reflected by the modulus of the spectrum, and of temporal aspects, which are reflected by the phase of the spectrum.

As represented dashed in figure 23, the post-lifter 140 can also be adaptive, the adaptation performed by the

module 144 pertaining jointly to the two post-lifters 15, 140. In this case, the post-lifter 55 of the decoder (figure 8) is adapted, like the post-lifter 140, as a function of parameters iLif which the adaptation module 144 provides to the mutliplexer 6 so that it includes them in the digital stream Φ. Typically, a few sets of coefficients γ_1, γ_2 are envisaged for the post-lifters 140 and 55, and the module 144 carries out an exhaustive test of these various sets of coefficients so as to retain the one which minimizes the discrepancy in the complex plane.

In the example represented in figure 24, the adaptation module 29 for the post-lifter 15 is identical to that of figure 1. Figure 24 shows a module 145 for estimating a masking curve allowing the module 29 to select, for the minimization of the discrepancy in terms of modulus, the harmonic frequency or frequencies which most exceed the masking curve calculated on the basis of the modulus spectrum $|X|$, as described above.

The post-lifter 140 of figure 24 is adapted separately by a module 146 which carries out the minimization of the discrepancies between the phase φ_X of the spectrum of the signal and the minimum phase $\varphi(k)$ calculated by the module 142 for one or more of the harmonics. Here, again, the harmonics selected for the calculation of the minimized phase discrepancy may be so as a function of the masking curve estimated by the module 145. The module 146 provides the output multiplexer 6 of the coder with the parameters iLif which represent the optimal post-lifter 140, so that they are used in the post-filter 55 of the decoder.

In the example illustrated by figure 25, the post-lifter 140 serving in the calculation of the minimum phases is not adaptive. The minimum phases $\varphi(k)$ calculated by the module 142 for the harmonics of the fundamental frequency are compared with the phases φ_X

of the spectrum of the audio signal, and the phase discrepancy forms the subject of a quantization by a module 148. The corresponding quantization indices $i\Delta\varphi$ are provided by the module 148 to the output multiplexer 6 of the coder.

In a decoder (figure 26) corresponding to a coder according to figure 25, a module 149 utilizes these quantization indices $i\Delta\varphi$ provided by the demultiplexer 45 to obtain the values of the quantized phase discrepancies, which are added by an adder 150 to the minimum phases $\varphi(k)$ calculated by the module 57 (the post-lifters 140 and 55 being identical). The phases provided by the adder 150 are then used by the module 54 which synthesizes the spectral lines of the harmonic component X_v.

The phase discrepancy quantized by the module 148, and which is used by the modules 149 and 150 of the decoder to correct the minimum phases $\varphi(k)$, can be of two kinds:

- it can represent, for each frequency of index i corresponding to a harmonic of order k of the fundamental frequency F_0, the difference between the phase $\varphi_X(i)$ of the spectrum of the signal at the frequency i and the minimum phase $\varphi(k)$ calculated by the module 142 for harmonic k;

- alternatively or cumulatively, this phase discrepancy can represent the variation of the phase φ_X of the spectrum over the width of one or more spectral peaks corresponding to harmonics of the signal, this variation relating to the minimum phase $\varphi(k)$ assigned to the peaks in question.

In both cases, the peak or peaks for which the phase discrepancy is quantized may be chosen as a function of the spectral energy represented by the upper envelope,

which is available to the coder and to the decoder, thereby enabling the decoder to determine that spectral line to which the discrepancies should be applied.

In the first case, the phase discrepancies may form the subject of a scalar quantization, or a vector quantization if they are grouped together for several peaks.

In the second case, the variation of the phase φ_X around the minimum phase $\varphi(k)$ over the width of a harmonic peak (determined by the width of the reference line used by the module 54), can be represented simply by the slope of a linear segment selected as being that which exhibits a minimum quadratic distance with the curve of the variation in phase of the spectrum over the width of the line, and possibly by a shift at the origin.

These slopes may form the subject of a scalar quantization, or a vector quantization if they are grouped together for several peaks.

The quantization of the phase variations over the harmonic peaks may pertain to the collection of harmonic frequencies. Another possibility is to quantize several slopes each obtained by averaging the slopes at the harmonics over one or more subbands of the spectrum. This averaging can be weighted so as to take account of the energies relating to the various harmonic frequencies, represented by the upper envelope.

The module 148 can also model the phase variation over the width of a peak by a more complex curve than a linear segment, for example a spline, whose parameters are quantized so as to be transmitted to the decoder.

Another possibility is to perform prior learning of

phase models at the harmonics, representative of the phase variations over the width of the peaks, which variations are observed in a corpus of reference signals. These models are held in a dictionary stored by the modules 148 and 149. The module 148 of the coder determines the indices $i\Delta\varphi$ corresponding to the addresses of the models closest to the phase variations in the neighborhood of the harmonic peaks considered, and the module 149 of the decoder recovers these models for the synthesis of the phase of the harmonic component.

CLAIMS

1. A method of coding an audio signal (x), in which a fundamental frequency (F_0) of the audio signal is estimated, a spectrum of the audio signal is determined through a transform in the frequency domain of a frame of the audio signal, and data for coding a harmonic component of the audio signal, comprising data representative of spectral amplitudes associated with frequencies which are multiples of the fundamental frequency, are included in a digital output stream (Φ), in which the spectral amplitude associated with one of said frequencies which are multiples of the fundamental frequency is a local maximum of the modulus of the spectrum in the neighborhood of said multiple frequency, and in which said data representative of spectral amplitudes associated with frequencies which are multiples of the fundamental frequency (F_0) are obtained by means of cepstral coefficients (cx_sup) calculated by transforming in the cepstral domain a compressed upper envelope (LX_sup) of the spectrum of the audio signal.

2. The method as claimed in claim 1, in which the compressed upper envelope (LX_sup) is determined by interpolation of said spectral amplitudes associated with the frequencies which are multiples of the fundamental frequency (F_0), with application of a spectral compression function.

3. The method as claimed in claim 2, in which the interpolation is performed between points whose abscissa is a frequency which is a multiple of the fundamental frequency (F_0) and whose ordinate is the spectral amplitude associated with said multiple frequency, compressed or uncompressed.

4. The method as claimed in any one of the preceiding claims, in which the transformation in the cepstral domain of the compressed upper envelope (LX_sup) is performed according to a nonlinear frequency scale.

5. The method as claimed in any one of the preceding claims, in which the cepstral coefficients (cx_sup) are quantized so as to form said data representative of the spectral amplitudes associated with the frequencies which are multiples of the fundamental frequency (F_0).

6. The method as claimed in claim 5, in which the quantization of the cepstral coefficients (cx_sup) pertains to a prediction residual for each of the cepstral coefficients.

7. The method as claimed in claim 6, in which the prediction residual for a cepstral coefficient is of the form $(cx[n,i] - \alpha(i)\cdot rcx\_q[n-1,i])/[2-\alpha(i)]$, where cx[n,i] designates a current value of said cepstral coefficient, rcx_q[n-1,i] designates a previous value of the quantized prediction residual, and $\alpha(i)$ designates a prediction coefficient.

8. The method as claimed in claim 6 or 7, in which different predictors are employed to determine the prediction residuals for at least two of the cepstral coefficients.

9. The method as claimed in any one of claims 5 to 8, in which the cepstral coefficients (cx_sup) are distributed into several cepstral subvectors quantized separately by a vector quantization pertaining to a prediction residual of the cepstral coefficients.

10. The method as claimed in any one of claims 5 to 9, in which the cepstral coefficients (cx_sup) are normalized before quantization, by modifying the cepstral coefficient of order 0 in such a way that the spectral amplitude associated with a frequency which is a multiple of the fundamental frequency (F_0) is represented exactly by the normalized cepstral coefficients.

11. The method as claimed in any one of claims 5 to 10, in which the cepstral coefficients (cx_sup) are transformed by liftering in the cepstral domain before being quantized.

12. The method as claimed in claim 11, in which the liftering is of the form $c_p(i) = [1 + \gamma_2^i - \gamma_1^i].c(i) - (\mu^i/_i)$, where $c_p(i)$ and $c(i)$ designate the cepstral coefficient of order $i>0$ respectively before and after liftering, γ_1 and γ_2 are coefficients lying between 0 and 1 and μ is a pre-emphasizing coefficient.

13. The method as claimed in claim 12, in which $\mu = (\gamma_2 - \gamma_1).c(1)$.

14. The method as claimed in any one of claims 11 to 13, in which a value of the modulus of the spectrum of the audio signal at at least one frequency which is a multiple of the fundamental frequency (F_0) is recalculated on the basis of the transformed and quantized cepstral coefficients (cx_sup_q), and said liftering is adapted in such a way as to minimize a discrepancy in modulus between the spectrum of the audio signal and at least one recalculated modulus value.

15. The method as claimed in any one of claims 11 to

13, in which a value of the modulus of the spectrum of the audio signal at at least one frequency which is a multiple of the fundamental frequency (F_0) is recalculated on the basis of the transformed and quantized cepstral coefficients (cx_sup_q), the cepstral coefficients are retransformed by liftering and smoothing in the cepstral domain, minimum phases ($\varphi(k)$) of the audio signal at frequencies which are multiples of the fundamental frequency are calculated on the basis of the retransformed cepstral coefficients (cxl[n]), and the liftering performed before the quantization is adapted in such a way as to minimize a deviation between the spectrum of the audio signal and at least one complex value whose modulus has a value recalculated for a frequency which is a multiple of the fundamental frequency and whose phase is given by the minimum phase calculated for said multiple frequency.

16. The method as claimed in claim 15, in which the lifterings performed before and after quantization are adapted jointly so as to minimize said discrepancy, and in which parameters (iLif) representative of the adapted liftering performed after quantization are included in the data for coding the harmonic component.

17. The method as claimed in any one of claims 14 to 16, in which the minimized discrepancy for the adaptation of the liftering relates to at least one frequency which is a multiple of the fundamental frequency (F_0), selected on the basis of the magnitude of the modulus of the spectrum in absolute value.

18. The method as claimed in any one of claims 14 to 16, in which a curve of spectral masking of the

audio signal is estimated by means of a psychoacoustic model, and the minimized discrepancy for the adaptation of the liftering relates to at least one frequency which is a multiple of the fundamental frequency (F_0), selected on the basis of the magnitude of the modulus of the spectrum in relation to the masking curve.

19. The method as claimed in claim 1, in which the spectrum of the audio signal and the cepstral coefficients (cx_sup) resulting from the transformation of the compressed upper envelope are determined for successive frames of N samples of the audio signal which exhibit mutual overlaps, and in which said data representative of spectral amplitudes associated with the frequencies which are multiples of the estimated fundamental frequency (f_0), obtained by means of the cepstral coefficients calculated by transforming the compressed upper envelope, are included in the digital output stream (Φ) for just one subset of the frames.

20. The method as claimed in claim 19, in which, for the frames which do not form part of said subset, data (icx[n-1/2]) for quantizing an error (ecx[n-1/2]) of interpolation of the cepstral coefficients resulting from the transformation of the compressed upper envelope (LX_sup) are included in the digital output stream (Φ).

21. The method as claimed in claim 19, in which, for the frames which do not form part of said subset, an optimal interpolator filter (128) is determined for the cepstral coefficients resulting from the transformation of the compressed upper envelope (LX_sup) and data (iP) representing said optimal interpolator filter are included in the digital

output stream (Φ).

22. An audio coder, comprising means for executing a method according to any one of the preceiding claims.

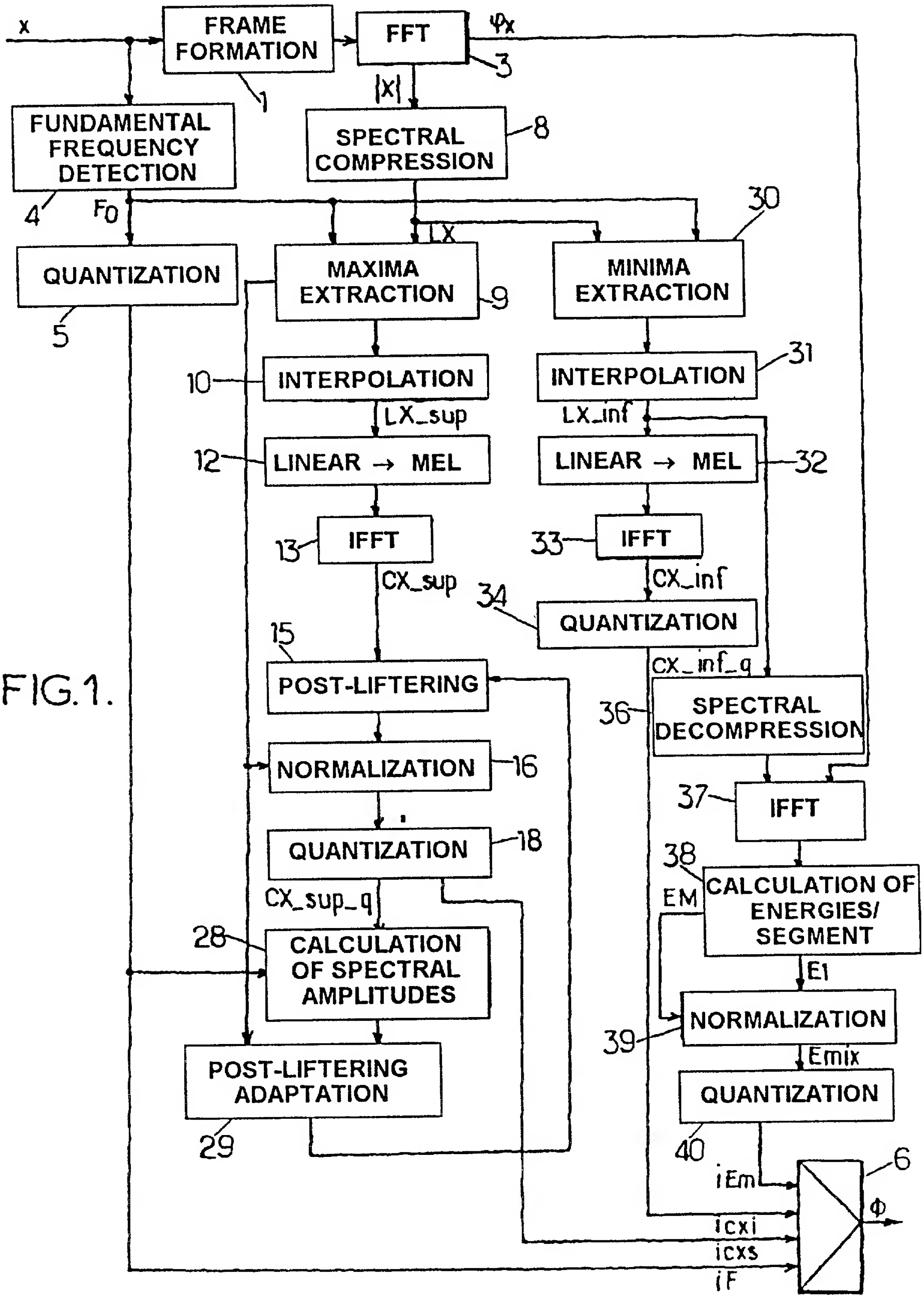

x
FRAME FORMATION
FFT
φx
3
1
|X|
FUNDAMENTAL FREQUENCY DETECTION
SPECTRAL COMPRESSION
8
4
F0
30
LX
QUANTIZATION
MAXIMA EXTRACTION
MINIMA EXTRACTION
9
5
INTERPOLATION
INTERPOLATION
10
31
LX_sup
LX_inf
LINEAR → MEL
LINEAR → MEL
12
32
IFFT
IFFT
13
33
CX_sup
CX_inf
34
QUANTIZATION
15
POST-LIFTERING
CX_inf_q
SPECTRAL DECOMPRESSION
36
NORMALIZATION
16
IFFT
37
QUANTIZATION
18
38
CALCULATION OF ENERGIES/ SEGMENT
EM
CX_sup_q
28
CALCULATION OF SPECTRAL AMPLITUDES
E1
NORMALIZATION
39
Emix
POST-LIFTERING ADAPTATION
QUANTIZATION
29
40
6
iEm
Φ
icxi
icxs
iF

FIG.1.

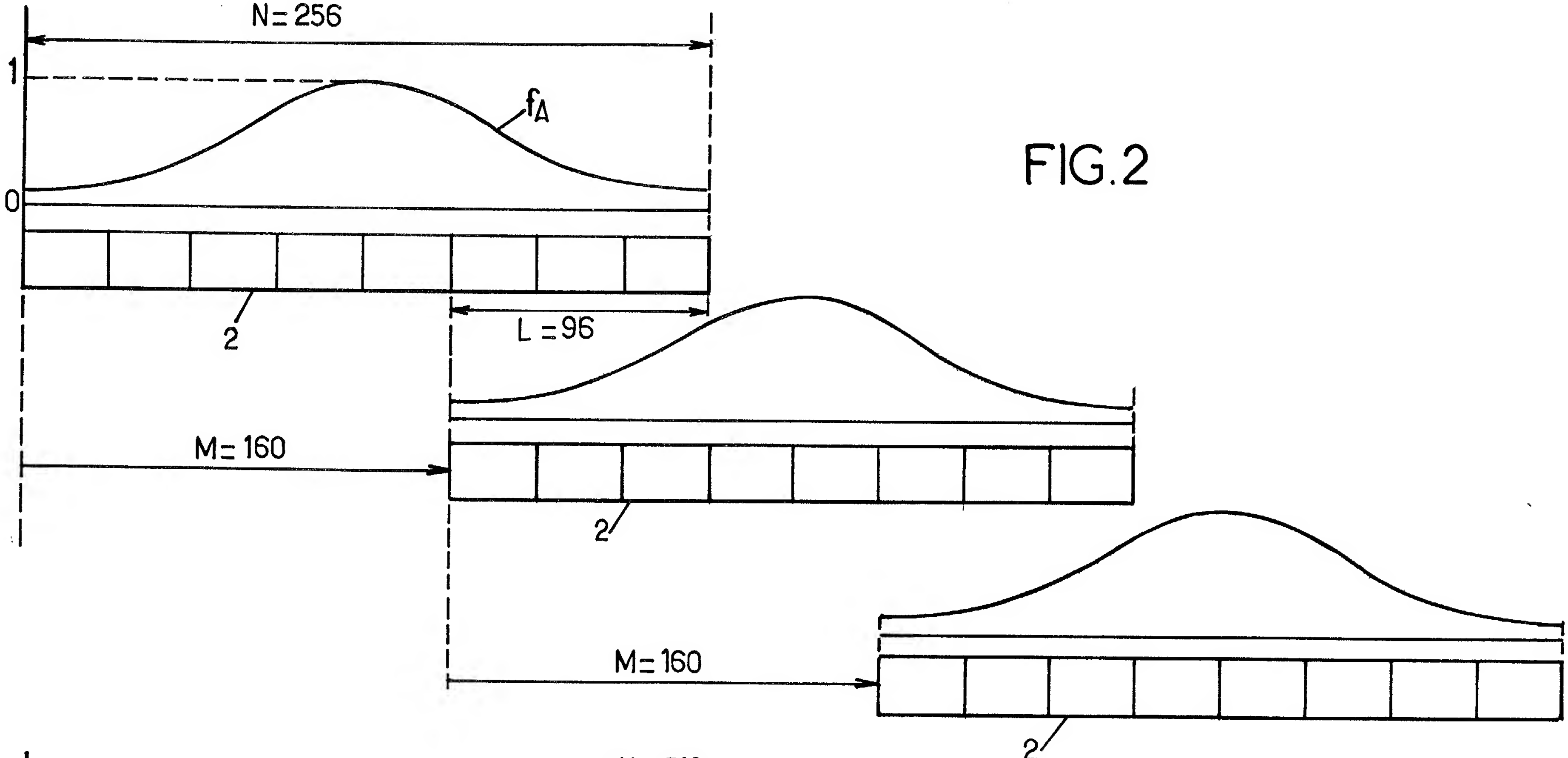


FIG.2

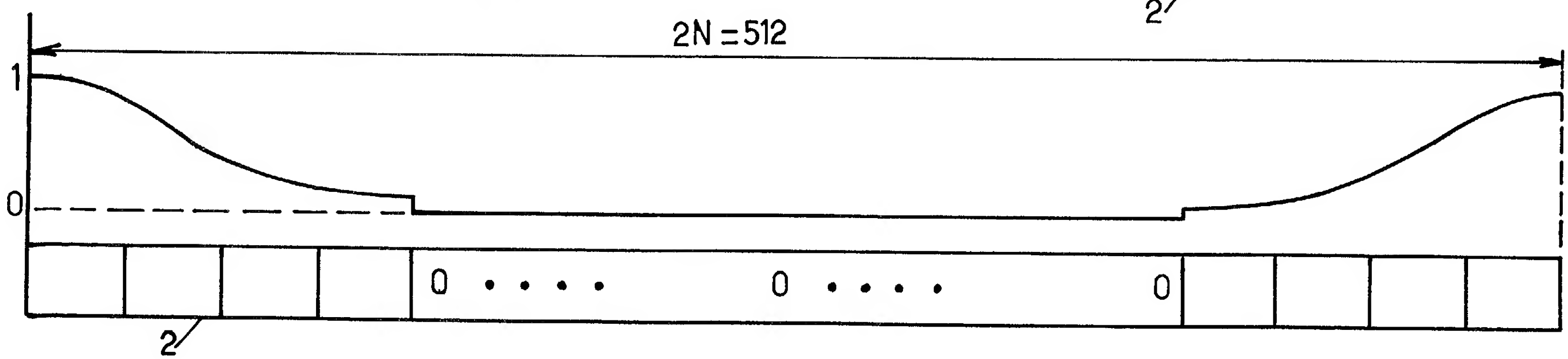


FIG.3.

FIG.4.

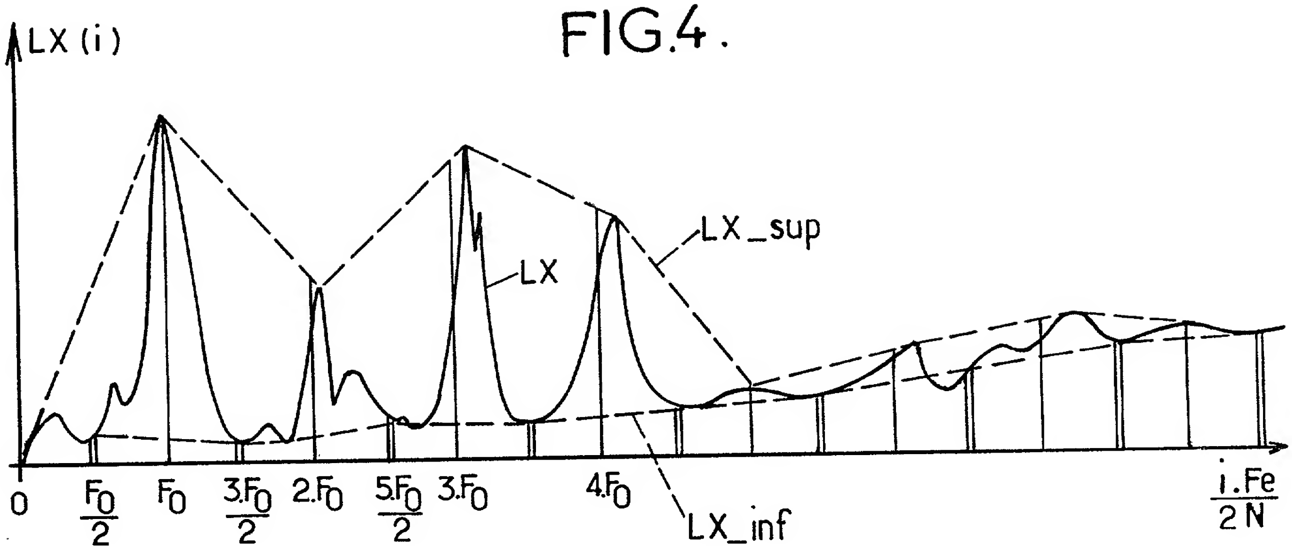

LX (i)
LX_sup
LX
LX_inf
0
$\frac{F_0}{2}$
F_0
$\frac{3.F_0}{2}$
$2.F_0$
$\frac{5.F_0}{2}$
$3.F_0$
$4.F_0$
$\frac{i.Fe}{2N}$

FIG.5.

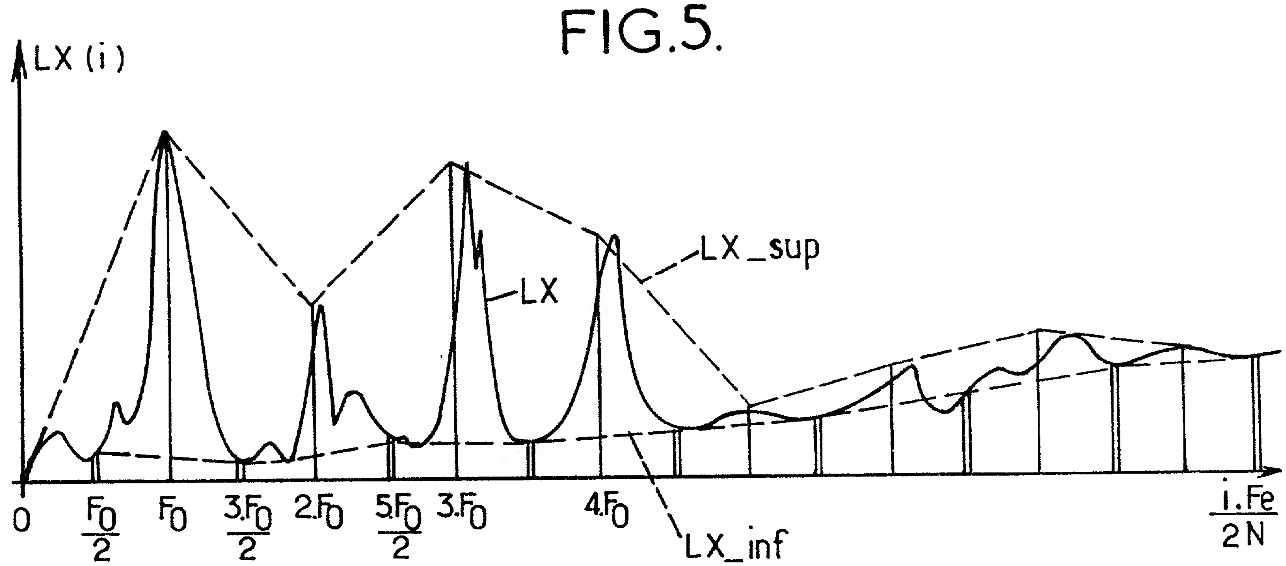

LX (i)
LX_sup
LX
LX_inf
0
$\frac{F_0}{2}$
F_0
$\frac{3.F_0}{2}$
$2.F_0$
$\frac{5.F_0}{2}$
$3.F_0$
$4.F_0$
$\frac{i.Fe}{2N}$

FIG.6.

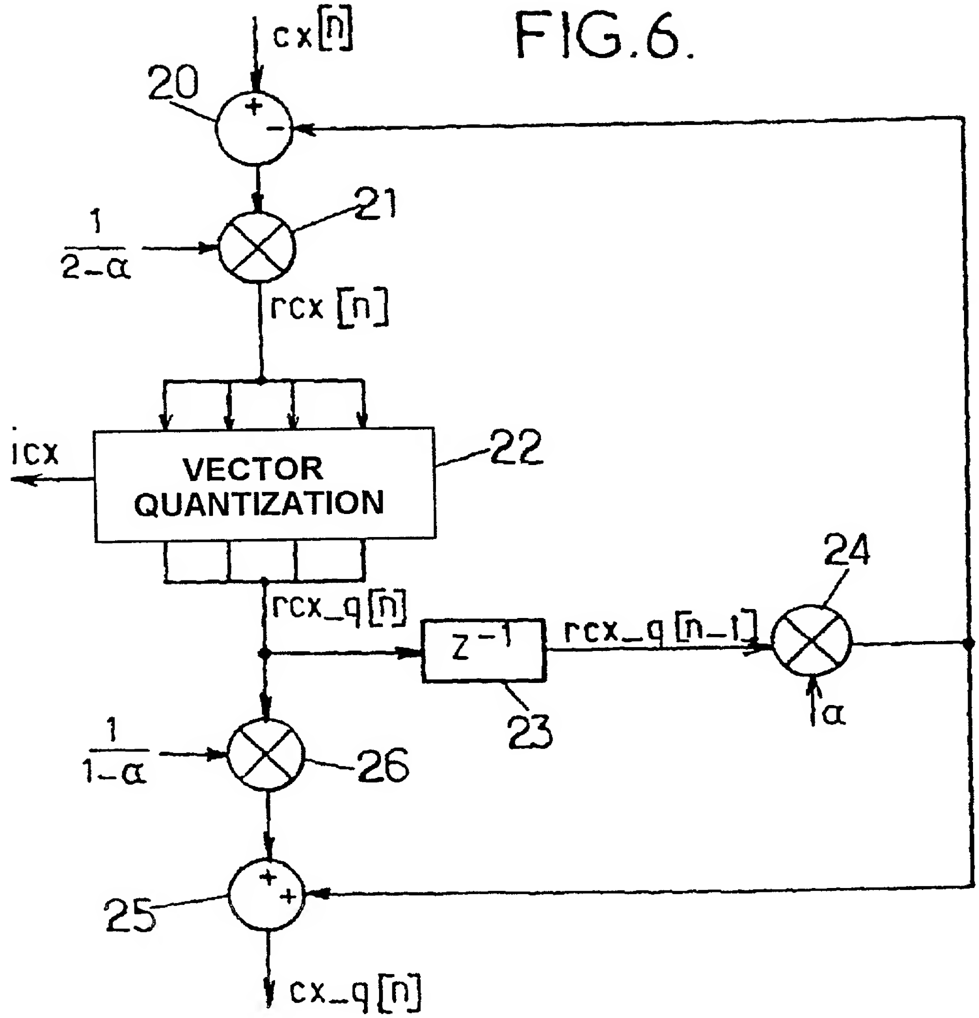
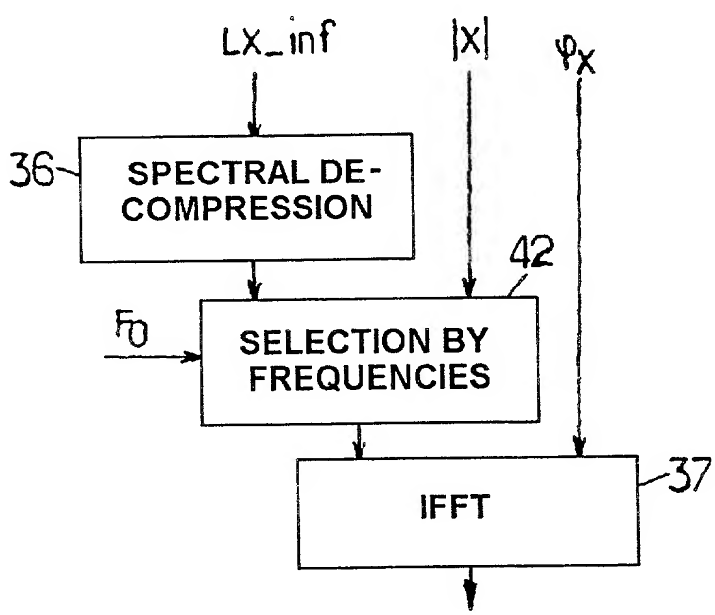

cx[n]
20
+
-
21
1
2-α
rcx [n]
icx
VECTOR
QUANTIZATION
22
24
rcx_q[n]
z-1
rcx_q [n-1]
α
23
1
1-α
26
25
+
+
cx_q[n]

FIG.7.


LX_inf
|X|
φX
36
SPECTRAL DE-
COMPRESSION
42
F0
SELECTION BY
FREQUENCIES
IFFT
37

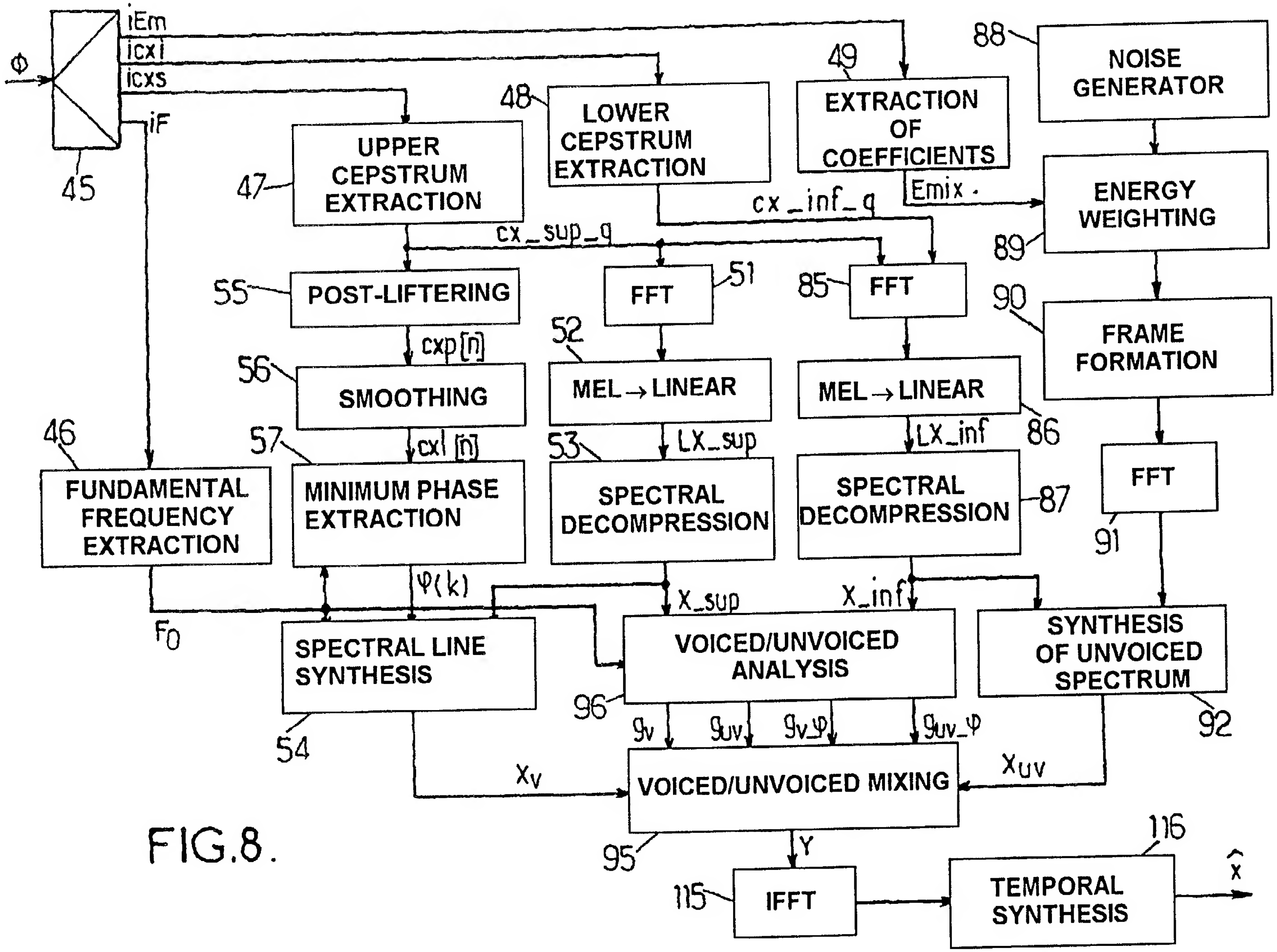
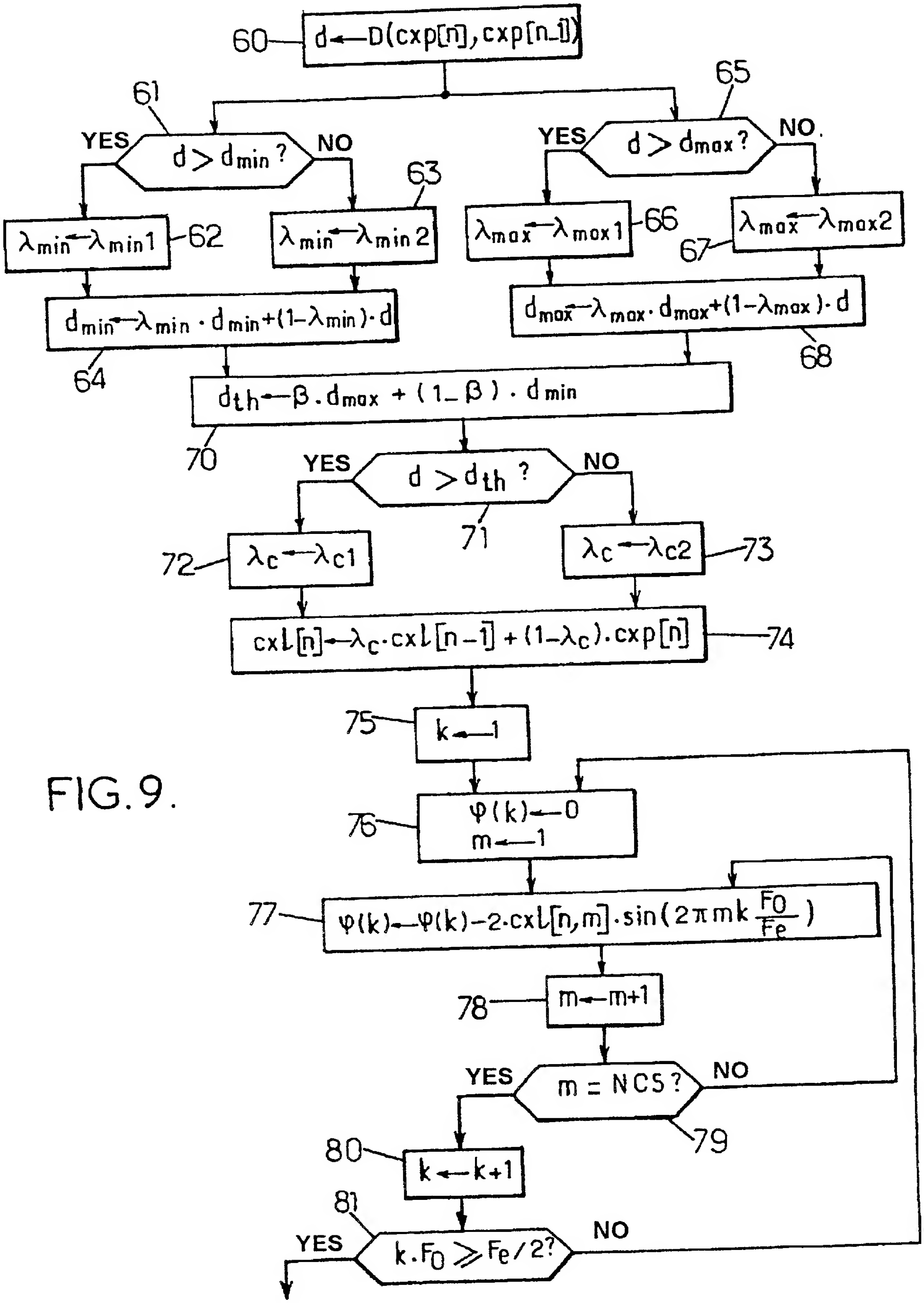

iEm
icxi
icxs
iF
45
46
FUNDAMENTAL FREQUENCY EXTRACTION
F0
47
UPPER CEPSTRUM EXTRACTION
48
LOWER CEPSTRUM EXTRACTION
49
EXTRACTION OF COEFFICIENTS
Emix
88
NOISE GENERATOR
89
ENERGY WEIGHTING
90
FRAME FORMATION
91
FFT
cx_sup_q
cx_inf_q
55
POST-LIFTERING
cxp[n]
56
SMOOTHING
cxl[n]
57
MINIMUM PHASE EXTRACTION
φ(k)
51
FFT
52
MEL→LINEAR
LX_sup
53
SPECTRAL DECOMPRESSION
X_sup
85
FFT
MEL→LINEAR
86
LX_inf
SPECTRAL DECOMPRESSION
87
X_inf
54
SPECTRAL LINE SYNTHESIS
Xv
96
VOICED/UNVOICED ANALYSIS
gv
guv
gvφ
guv_φ
92
SYNTHESIS OF UNVOICED SPECTRUM
Xuv
95
VOICED/UNVOICED MIXING
Y
115
IFFT
116
TEMPORAL SYNTHESIS
x̂

FIG.8.



60
d ← D(cxp[n], cxp[n-1])
61
d > dmin ?
YES
NO
62
λmin ← λmin1
63
λmin ← λmin2
64
dmin ← λmin · dmin + (1−λmin) · d
65
d > dmax ?
YES
NO
66
λmax ← λmax1
67
λmax ← λmax2
68
dmax ← λmax · dmax + (1−λmax) · d
70
dth ← β . dmax + (1−β) . dmin
71
d > dth ?
YES
NO
72
λc ← λc1
73
λc ← λc2
74
cxL[n] ← λc · cxL[n−1] + (1−λc) . cxp[n]
75
k ← 1
76
φ(k) ← 0
m ← 1
77
φ(k) ← φ(k) − 2 · cxL[n,m] · sin(2πmk F0/Fe)
78
m ← m+1
79
m = NCS ?
YES
NO
80
k ← k+1
81
k.F0 ≥ Fe/2?
YES
NO

FIG. 9.

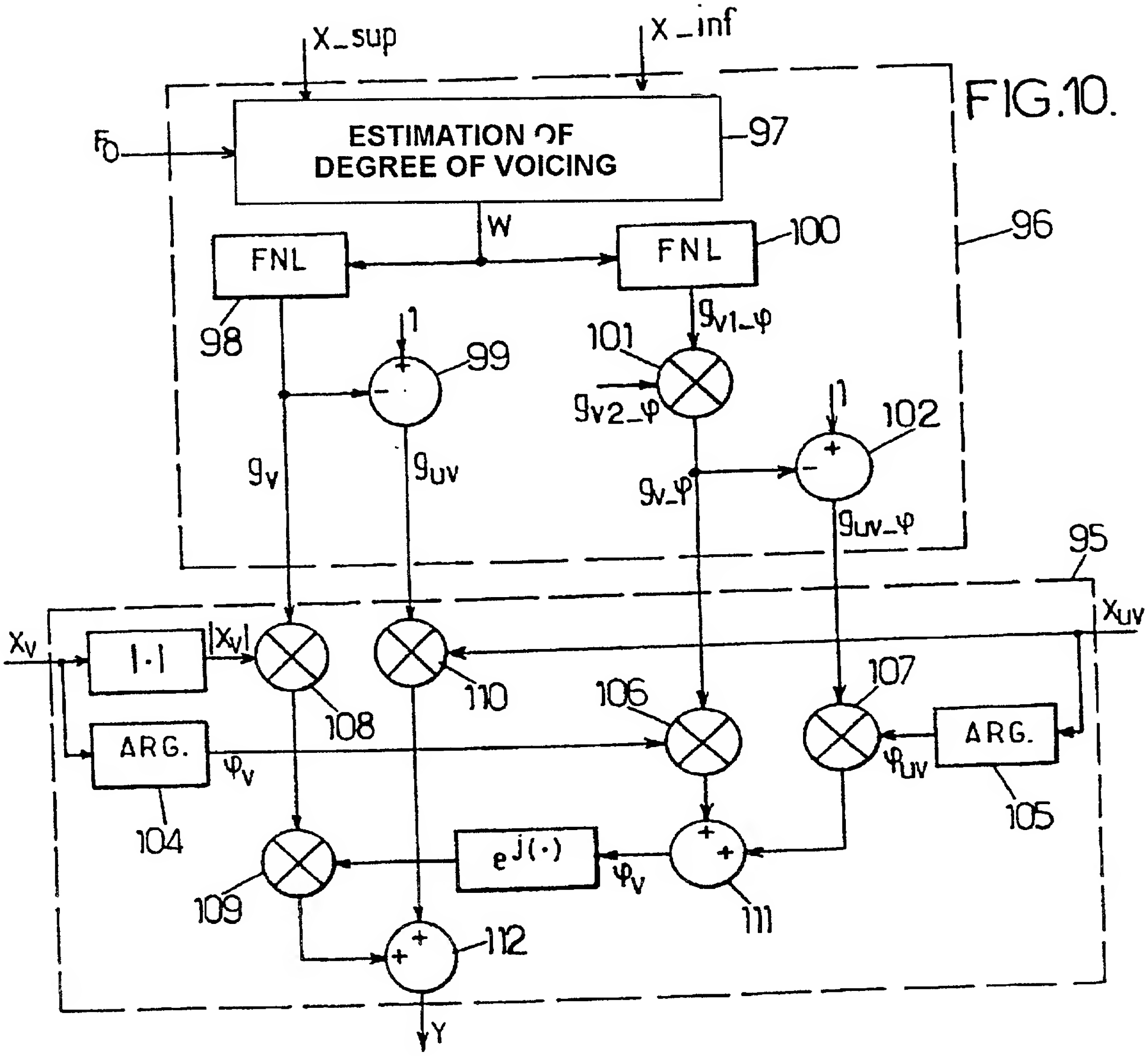


FIG.10.

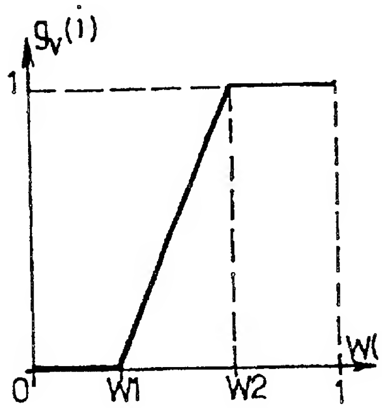


FIG.11

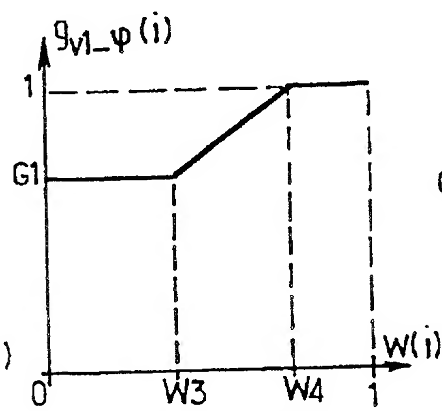


FIG.12.

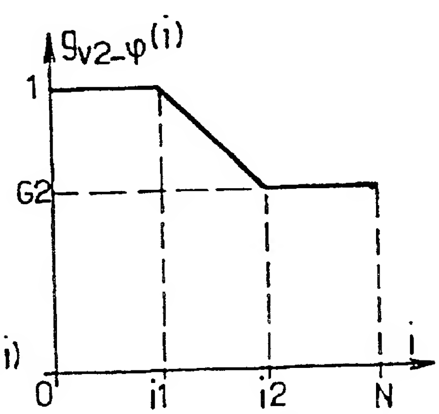


FIG.13.

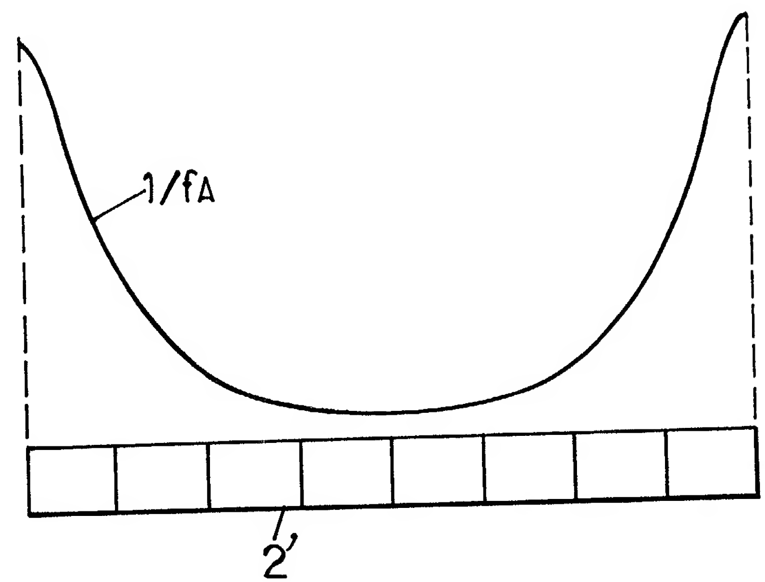


FIG.14.

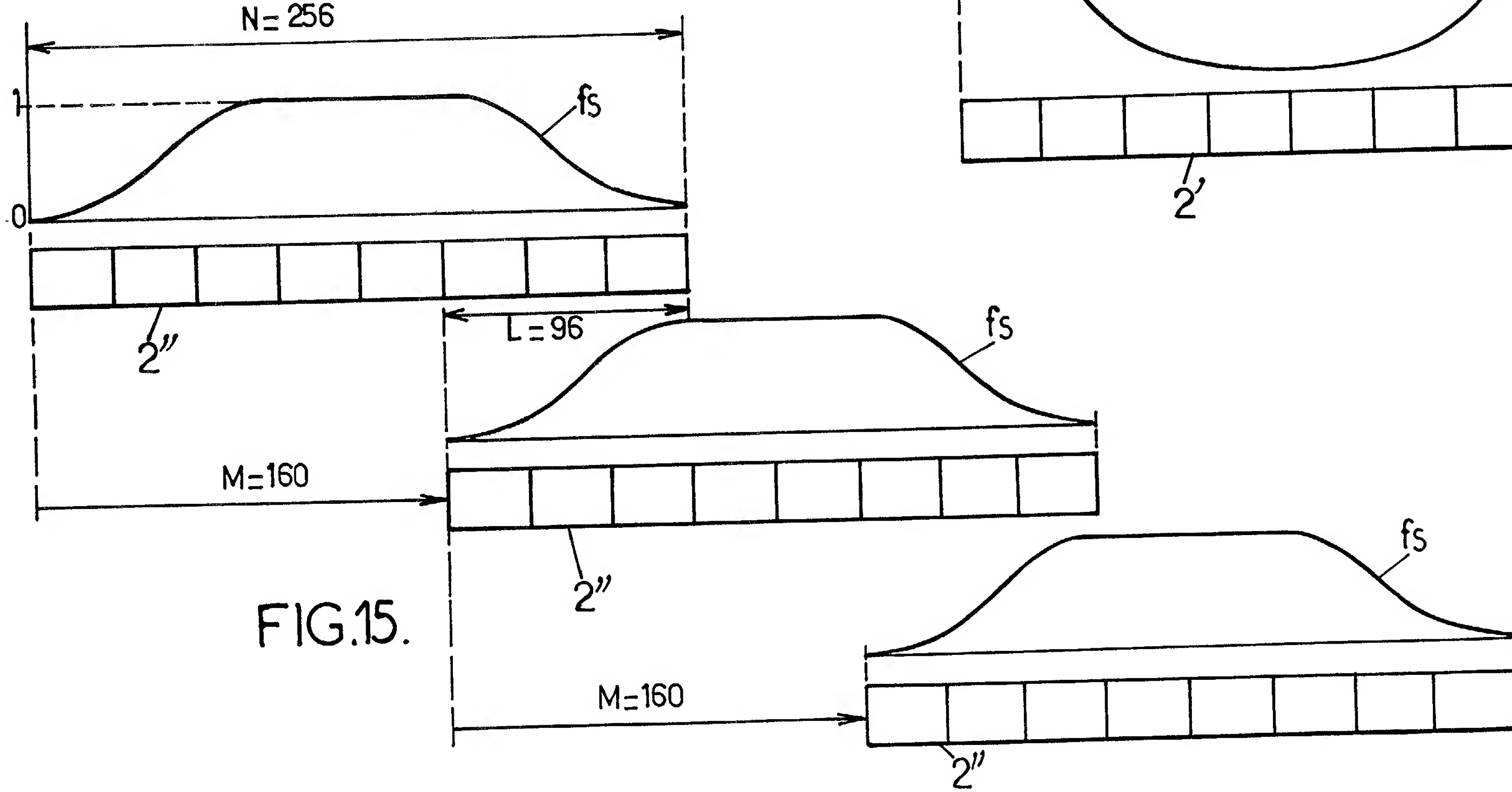


FIG.15.

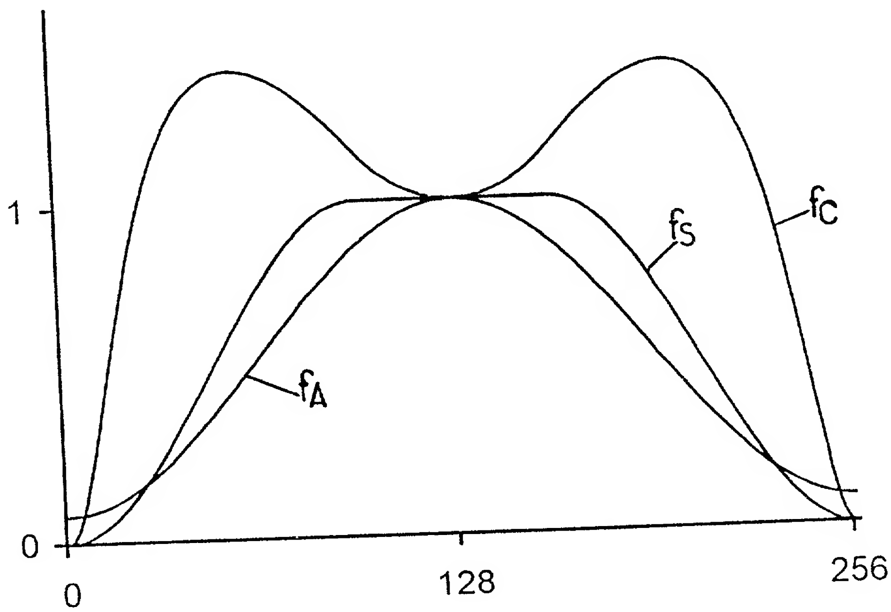


FIG.16.

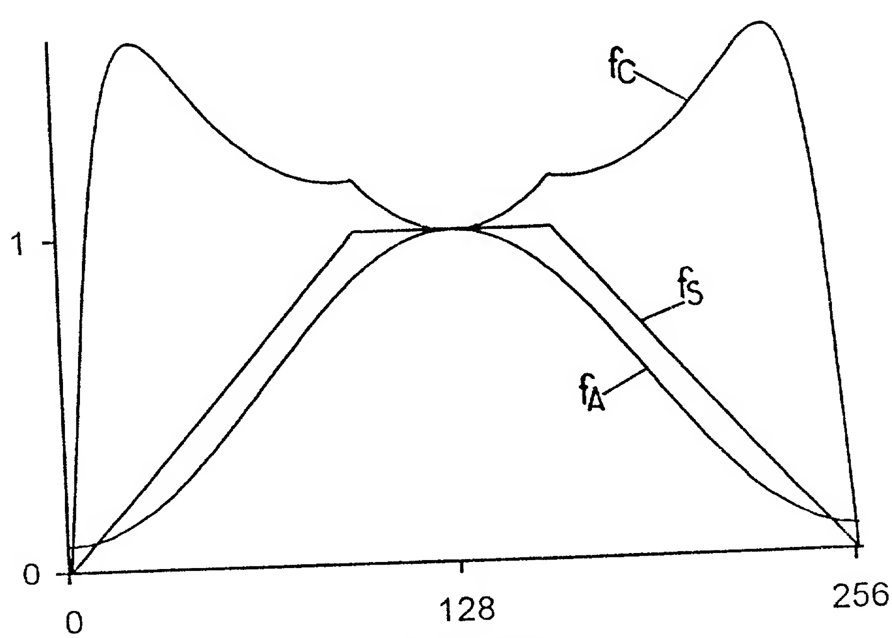


FIG.17.

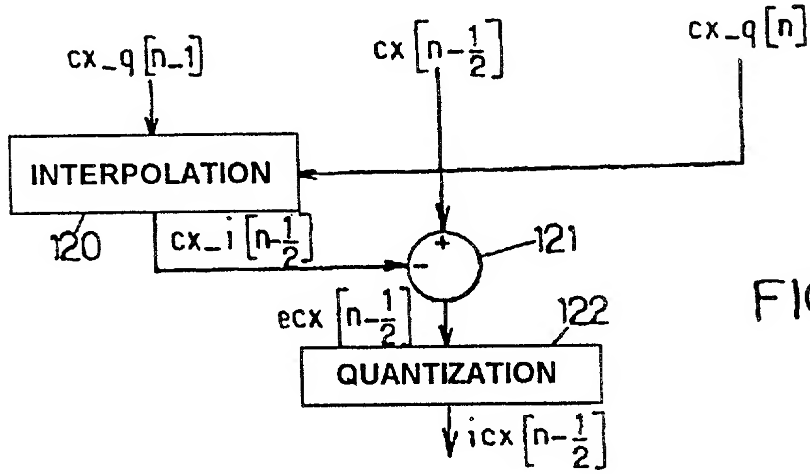


FIG.18.

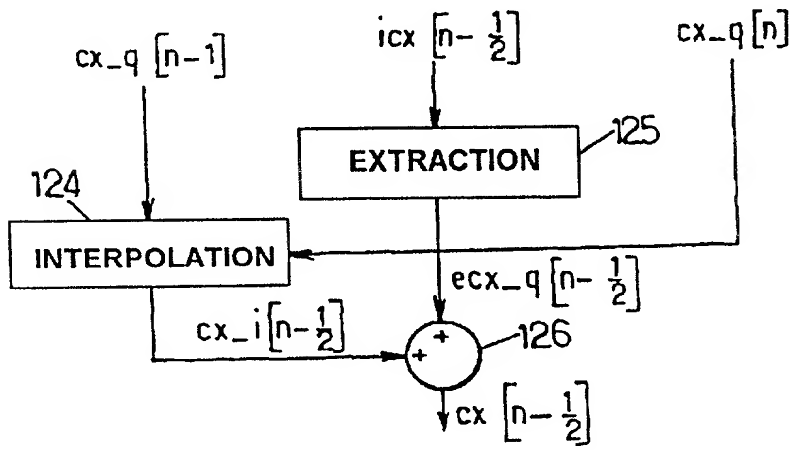


FIG.19.

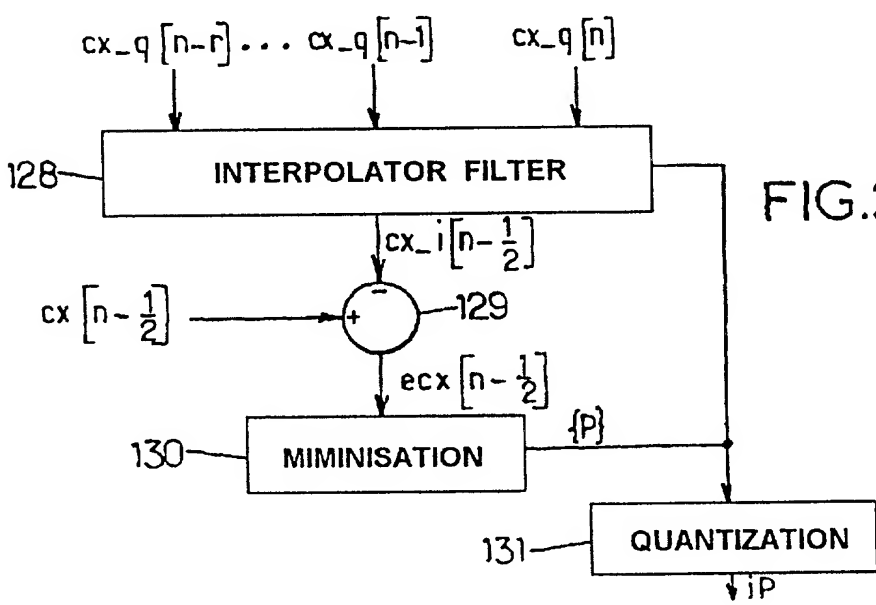


FIG.20.

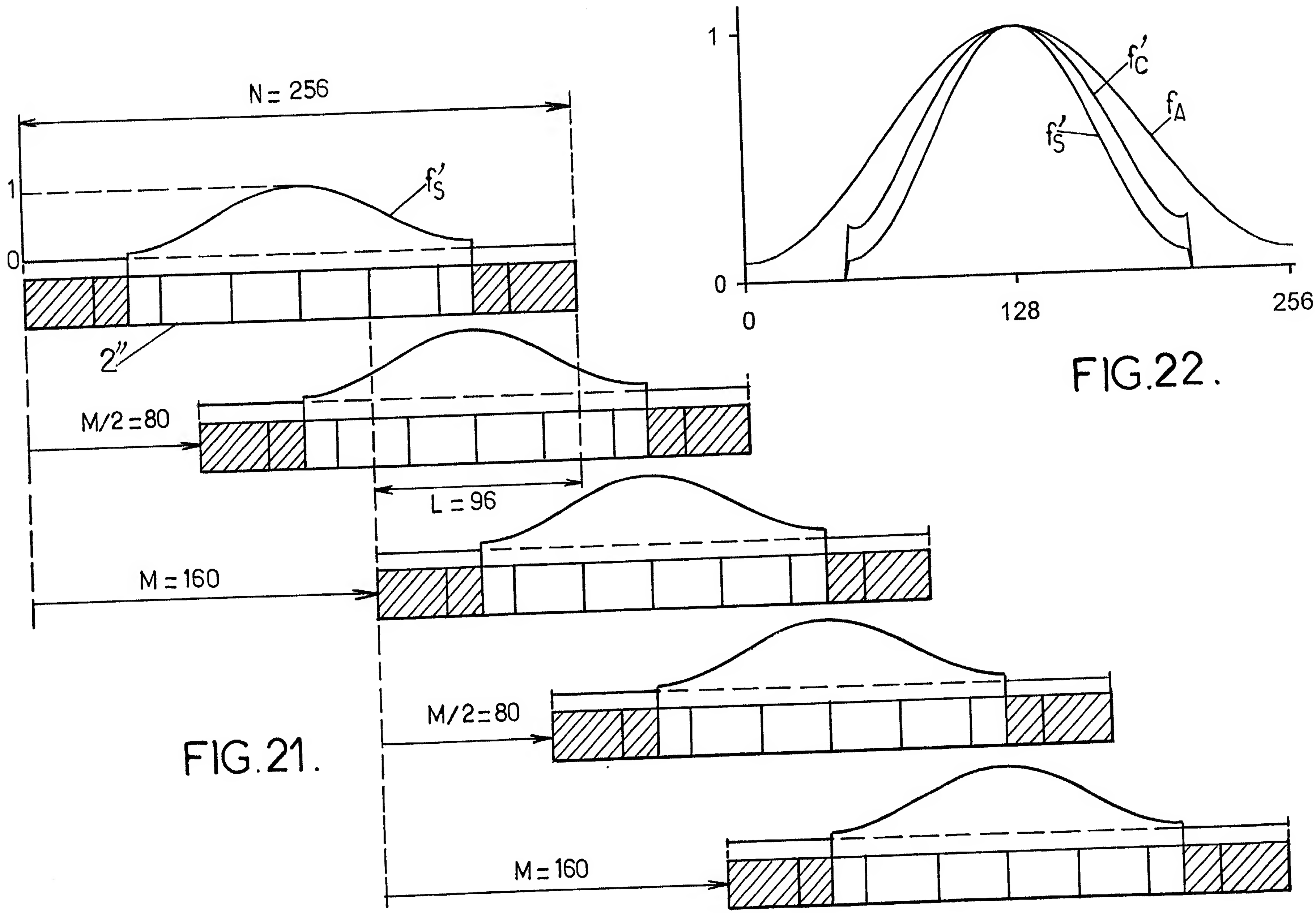


FIG.21.

FIG.22.

10/019965

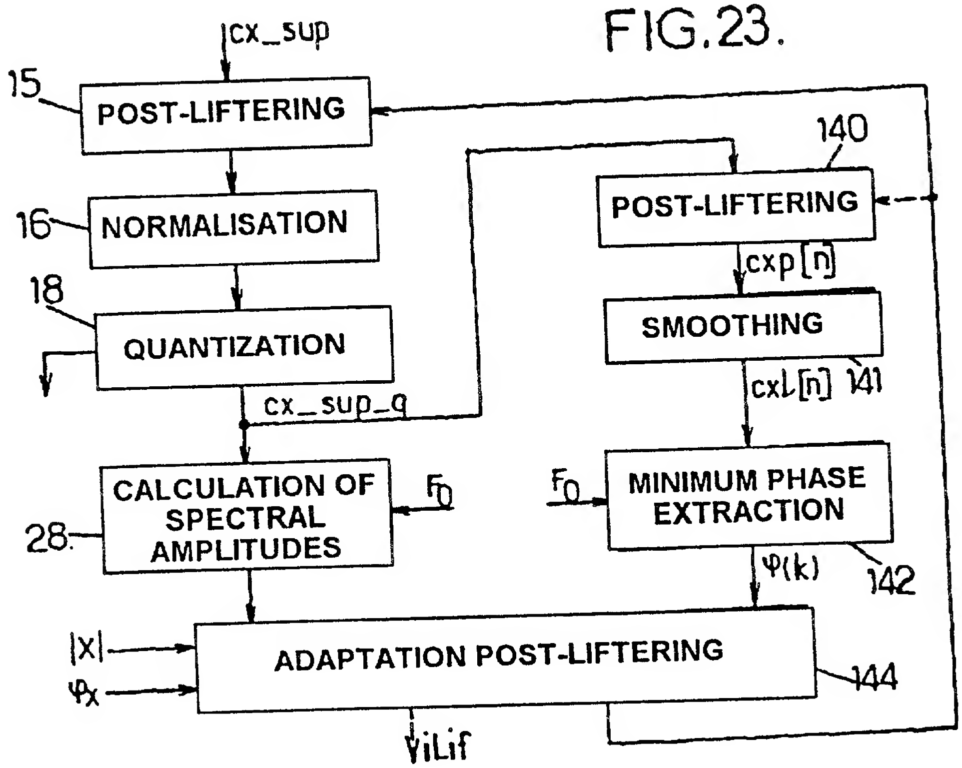


FIG.23.

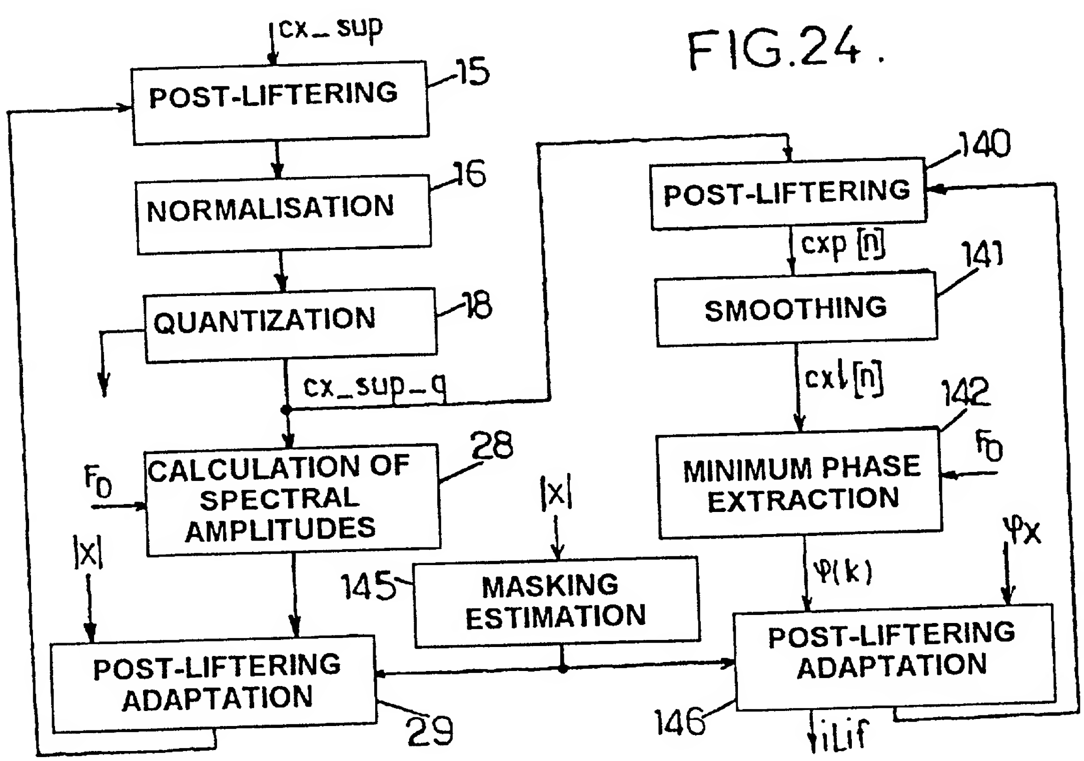


FIG.24.

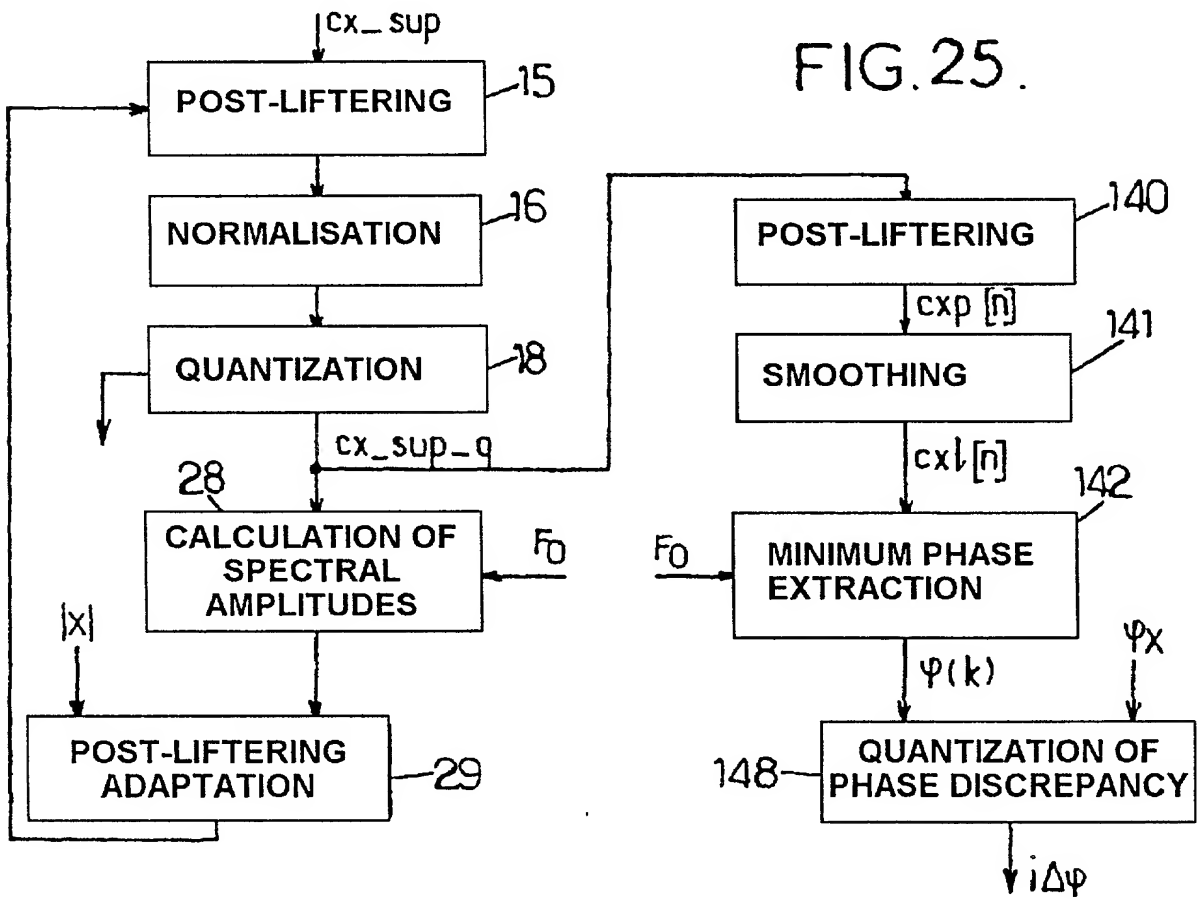

cx_sup
POST-LIFTERING
15
NORMALISATION
16
QUANTIZATION
18
cx_sup_q
28
CALCULATION OF SPECTRAL AMPLITUDES
F0
|X|
POST-LIFTERING ADAPTATION
29
POST-LIFTERING
140
cxp[n]
SMOOTHING
141
cxl[n]
142
F0
MINIMUM PHASE EXTRACTION
φ(k)
φX
148
QUANTIZATION OF PHASE DISCREPANCY
iΔφ

FIG. 25.

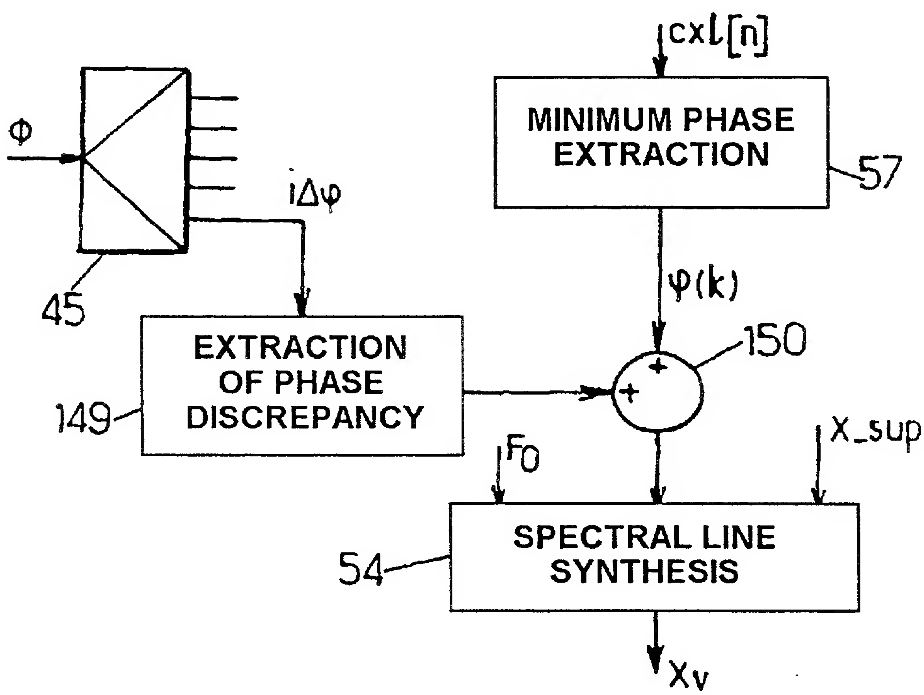

cxl[n]
MINIMUM PHASE EXTRACTION
57
Φ
iΔφ
45
φ(k)
150
+
+
EXTRACTION OF PHASE DISCREPANCY
149
F0
X_sup
54
SPECTRAL LINE SYNTHESIS
Xv

FIG. 26.

DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below, next to my name.

I believe I am the original, first, and sole inventor (if only one name is listed below) or an original, first, and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled
AUDIO CODING WITH HARMONIC COMPONENTS

the specification of which

X	is attached hereto.
	was filed on as
	United States Application Number ______
	Or PCT International Application Number ______
	And was amended on ______ (if applicable)

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above. I do not know and do not believe that the claimed invention was ever known or used in the United States of America before my invention thereof, or patented or described in any printed publication in any country before my invention thereof or more than one year prior to this application, that the same was not in public use or on sale in the United States of America more than one year prior to this application, and that the invention has not been patented or made the subject of an inventor's certificate Issued before the date of this application in any country foreign to the United States of America on an application filed by me or my legal representatives or assigns more than twelve months (for a utility patent application) or six months (for a design patent application) prior to this application.

I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal Regulations, Section 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d), of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Express Mail No. EL857974142US

Prior Foreign Application(s):

			Priority Claimed	
99 08634	FRANCE	05/07/1999	X	
Number	(Country)	(Day/Month/Year Filed)	Yes	No
Number	(Country)	(Day/Month/Year Filed)	Yes	No

I hereby claim the benefit under title 35, United States Code, Section 119(e) of the United States provisional application(s) listed below:

(Application Number)	(Filing Date)
(Application Number)	(Filing Date)

I hereby claim the benefit under Title 35, United States Code, Section 120 of any United States application(s) listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior United States application in the manner provided by the first paragraph of Title 35, United States Code, Section 112, I acknowledge the duty to disclose all information known to me to be material to patentability as defined in Title 37, Code of Federal regulations, Section 1.56 which became available between the filing date of the prior application and the national or PCT International filing date of this application:

PCT/FR00/01908	04/07/2000	
(Application Number)	Filing Date	(Status-patented, pending, abandoned)

I hereby appoint Timothy N. Trop, Reg. No. 28,994; Fred G. Pruner, Jr., Reg. No. 40,779, Dan C. Hu, Reg. No. 40,025 and Ruben S. Bains, Reg. No. 46,532; my patent attorneys, of TROP, PRUNER & HU, P.C., with offices located at 8554 Katy Freeway, Ste. 100, Houston, TX 77024, telephone (713) 468-8880, my patent attorneys; with full power of substitution and revocation, to prosecute this application and to transact all business in the Patent and Trademark Office connected herewith. Customer No. 21906

Send correspondence to Dan C. Hu, TROP, PRUNER & HU, P.C., 8554 Katy Freeway, Ste. 100, Houston, TX 77024 and direct telephone calls to Dan C. Hu, (713) 468-8880.

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

1-00

2-00

Full Name of Sole/First Inventor: François CAPMAN	
Inventor's Signature: [signature]	Date: 05 Dec 2001
Residence: 47 rue des Etats Généraux, 78000 VERSAILLES, FRANCE 1-12X	Citizenship: French
Post Office Address: THE SAME AS ABOVE	
Full Name of Second/Joint Inventor: Carlo MURGIA	
Inventor's Signature: X Carlo Murgia	Date: 05 Dec 2001
Residence: 458 Saint Vincent, IRVINE, CA CA 92618, U.S.A.	Citizenship: Italian
Post Office Address: THE SAME AS ABOVE	
Full Name of Third/Joint Inventor:	
Inventor's Signature:	Date:
Residence:	Citizenship:
Post Office Address:	
Full Name of fourth/Joint Inventor:	
Inventor's Signature:	Date:
Residence:	Citizenship:
Post Office Address:	